RL-TR-94-23
Final Technical Report
April 1994

PROTOTYPE RULE-BASED RELIABILITY ANALYSIS FOR VLSI CIRCUIT DESIGN

University of Illinois at Urbana-Champaign

Sung-Mo Kang, Yusuf Leblebici, Weishi Sun,
and Chin-Chi Teng

APPROVED FOR PUBLIC RELEASE; DISTRIBUTION UNLIMITED.

Rome Laboratory
Air Force Materiel Command
Griffiss Air Force Base, New York

This report has been reviewed by the Rome Laboratory Public Affairs Office (PA) and is releasable to the National Technical Information Service (NTIS). At NTIS it will be releasable to the general public, including foreign nations.

RL-TR-94-23 has been reviewed and is approved for publication.

APPROVED: Christopher J. Flynn

CHRISTOPHER J. FLYNN
Project Engineer

FOR THE COMMANDER:

HARVEY D. DAHLJELM, Colonel, USAF
Director, Reliability
Electromagnetics & Reliability Directorate

If your address has changed or if you wish to be removed from the Rome Laboratory mailing list, or if the addressee is no longer employed by your organization, please notify RL (ERDD) Griffiss AFB NY 13441. This will assist us in maintaining a current mailing list.

Do not return copies of this report unless contractual obligations or notices on a specific document require that it be returned.

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. AGENCY USE ONLY (Leave Blank)	2. REPORT DATE	3. REPORT TYPE AND DATES COVERED
	April 1994	Final Oct 92 - Oct 93

4. TITLE AND SUBTITLE
PROTOTYPE RULE-BASED RELIABILITY ANALYSIS FOR VLSI CIRCUIT DESIGN

5. FUNDING NUMBERS
C - F30602-92-C-0069
PE - 62702F
PR - 2338
TA - 01
WU - PD

6. AUTHOR(S)
Sung-Mo Kang, Yusuf Leblebici, Weishi Sun, and Chin-Chi Teng

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
University of Illinois at Urbana-Champaign
Coordinated Science Laboratory
1308 West Main Street
Urbana IL 61801-2307

8. PERFORMING ORGANIZATION REPORT NUMBER
N/A

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)
Rome Laboratory (ERDD)
525 Brooks Road
Griffiss AFB NY 13441-4505

10. SPONSORING/MONITORING AGENCY REPORT NUMBER
RL-TR-94-23

11. SUPPLEMENTARY NOTES
Rome Laboratory Project Engineer: Christopher J. Flynn/ERDD/(315) 330-4102

12a. DISTRIBUTION/AVAILABILITY STATEMENT
Approved for public release; distribution unlimited.

12b. DISTRIBUTION CODE

13. ABSTRACT (Maximum 200 words)

This report describes the development and application of parametric and geometry-based macro-models of hot-carrier induced dynamic degradation in MOS VLSI circuits. Previously, a simulation based approach has been used for reliability analysis, but this is inefficient for reliability assessment of very large scale integrated circuits. Geometry-based macro-models for hot-carrier reliability estimation have been developed. The macro-models express hot-carrier damage as functions of designable parameters such as transistor size (W), output loading capacitance (C_1) and the input signal slew rate (a). A prototype rule-based reliability diagnosis tool, iRULE, has been developed. This tool uses the macro-models for designing hot-carrier resistant circuits without the need for transient reliability simulations. This provides the ability to analyze very large circuits with more than one million transistors on a workstation in a short amount of time. This report also describes a fast timing reliability simulation tool, ILLIADS-R, that can accurately estimate hot-carrier degradation while providing several orders of magnitude speed up over traditional transistor-level circuit simulators.

14. SUBJECT TERMS	15. NUMBER OF PAGES
Reliability, hot-carrier degradation, VLSI CMOS circuits, simulation	88
	16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT	18. SECURITY CLASSIFICATION OF THIS PAGE	19. SECURITY CLASSIFICATION OF ABSTRACT	20. LIMITATION OF ABSTRACT
UNCLASSIFIED	UNCLASSIFIED	UNCLASSIFIED	UL

FINAL REPORT EVALUATION

Contract F30602-92-C-0069

PROTOTYPE RULED-BASED RELIABILITY ANALYSIS FOR VLSI CIRCUIT DESIGN

The growth in the use of custom, semi-custom, and application specific ICs (ASICs) in AF systems has brought several changes to the reliability qualification process. This occurred because the costs associated with reliability characterization has become a major part of the device cost when there are only a few hundred or thousand of such devices produced. Further, the cost of a second or third pass at the design/fabricate/test/evaluate reliability cycle is very time consuming as well as expensive.

Analyzing an IC design for susceptibility to life limiting failure mechanisms is one approach to enhance the probability that at the end of the first pass of the design/fabricate/test/evaluate reliability cycle, the device will have an acceptable reliability.

This report describes the development and application of parametric and geometry-based macro-models of hot-carrier induced dynamic degradation in MOS VLSI circuits. The macro-models express hot-carrier damage as functions of designable parameters such as transistor size (W), output loading capacitance (C_l) and the input signal slew rate (a). A prototype rule-based reliability diagnosis tool, iRULE, has been developed. This tool uses the macro-models for designing hot-carrier resistant circuits without the need for transient reliability simulations. This provides the ability to analyze very large circuits with more than one million transistors on a workstation in a short amount of time.

This report also describes a fast timing reliability simulation tool, ILLIADS-R, that can accurately estimate hot-carrier degradation while providing several orders of magnitude speed up over traditional transistor-level circuit simulators.

Christopher J. Flynn

CHRISTOPHER J. FLYNN
Project Engineer
RL/ERDD

Accesion For	
NTIS CRA&I	☑
DTIC TAB	☐
Unannounced	☐
Justification	
By	
Distribution /	
Availability Codes	
Dist	Avail and / or Special
A-1	

PROTOTYPE RULE-BASED RELIABILITY ANALYSIS FOR VLSI CIRCUIT DESIGN

Final Report

Contract F30602-92-C-0069 (Task N-3-0016)

ABSTRACT

This report describes the accomplishments during the contract period (October 13, 1992, to October 12, 1993) on the computer-aided analysis of CMOS device and circuit degradation due to hot-carrier effects. The task involved four subtasks: (1) Parametric macromodeling of hot-carrier-induced dynamic degradation in MOS VLSI circuits; (2) The development of the rule-based circuit reliability diagnosis program iRULE; (3) Tester chip design to verify the geometry-based macromodels; (4) Fast timing simulation tool ILLIADS-R to estimate hot-carrier-induced device and circuit degradation.

In order to estimate the hot-carrier reliability for circuits of over one million transistors in short time, we have developed new macromodels for hot-carrier-damaged CMOS logic circuits. In these macromodels, the hot-carrier damage can be expressed as functions of designable parameters such as transistor size (W), output loading capacitance (C_L) and the input signal slew rate (a). Based on these macromodels, a reliability rule checking program iRULE has been developed. iRULE uses a table-lookup method to quickly pinpoint the critical transistors in the circuit. iRULE has also been successfully technology transferred to AT&T Bell Labs to estimate the circuit hot-carrier lifetime.

The fast timing reliability simulation tool ILLIADS-R is also used for circuit hot-carrier reliability estimation. A hierarchical reliability evaluation system can be formalized to achieve both quick diagnosis and accurate reliability analysis of very large scale integrated circuits.

1 Introduction

Hot-carrier induced damages in MOS transistors cause reliability problems in VLSI circuits. The conventional approach of post-fabrication reliability qualification has become increasingly costly, causing a significant lengthening of the design cycle, especially when design changes are mandated due to unforeseen reliability problems. It is being recognized that important reliability issues have to be addressed rigorously in the design phase long before the post-manufacture reliability qualification. The development and use of accurate reliability simulation tools are therefore crucial for early assessment and improvement of circuit reliability: the long-term reliability of the circuit estimated through simulation can be compared with predetermined reliability specifications or limits. If the predicted reliability does not satisfy the requirements, appropriate design modifications should be made to improve the resistance of the devices to degradation. This cycle may be repeated several times until the simulated reliability of the circuit meets the desired specifications. The ultimate goal of CAD for reliability is to develop user-friendly and useful software tools that can be readily integrated into existing CAD frameworks and environments.

Previous efforts toward building accurate reliability estimators have focused on developing simulation-based tools [1, 2, 3, 4, 5, 6]. Reliability simulation is, however, inherently slow, particularly since the transient electrical analysis must incorporate the aging of the device parameters during simulation. Much more efficient CAD tools—such as reliability diagnosis programs—are required in order to solve this bottleneck problem. The reliability simulation of such large circuits would simply take too long to be practical. The case of the layout rule checker is an illustrative parallel: without such a program, layout designers would have been forced to resort to process simulation tools to detect potential shorts and opens in interconnects, etc. Layout rule checking tools based upon worst-case analysis have been successfully used for the layout design of VLSI circuits containing millions of transistors.

The strong push toward increasingly shorter design cycles without sacrificing reliability and performance mandates the development of reliability diagnosis tools. In particular, a fast reliability estimation tool and a design-for-reliability aid which can provide the designer with quick feedback are very much needed.

In order to meet such demands, a rule-based circuit reliability diagnosis program, iRULE, has been developed. Currently, iRULE is capable of quickly estimating hot-carrier induced device degradation in CMOS digital circuits, even for large circuits consisting of more than one million transistors. iRULE avoids simulation, but instead employs netlist- and geometry-based rules to determine the potentially critical transistors in the given circuit, and a macro-model (look-up) table containing hot-carrier degradation values as a function of circuit parameters. The table entries can be generated from the macro-models developed in [7, 8] or through experimental measurements and used to identify critical transistors. Since iRULE does not rely on transient simulation, hot-carrier reliability diagnosis of CMOS digital circuits can be carried out very efficiently.

iRULE is based on a set of geometry-based design rules to improve long-term reliability that can be readily applied to device and circuit design. The extent of hot-carrier damage that each nMOS transistor experiences during dynamic circuit operation is determined primarily by its terminal voltage waveforms, and it is influenced by such parameters as gate voltage rise time, drain voltage fall time, channel width, channel length, and input signal frequency. The challenge in developing reliability-oriented design rules is to identify a set of relevant device and circuit parameters, and to express the hot-carrier-related dynamic degradation of nMOS transistors under circuit operating conditions as a simple function of these design parameters. A parametric reliability measure is developed for estimating the hot-carrier induced degradation based on the layout geometry. The macro-model derivation is presented in this report. It is based on the well-known interface trap generation model [9]

and is consistent with degradation models used in most reliability simulators [1, 3, 4]. The analysis results obtained by using the macromodel for various simple circuit configurations agree with recently published experimental findings [10, 11].

A fast timing reliability simulation tool ILLIADS-R is also introduced. ILLIADS-R can provide a computational speed up of several orders of magnitude, while achieving an accuracy comparable to that of transistor-level circuit simulators. ILLIADS-R uses a newly developed simple empirical damaged MOSFET model to estimate hot-carrier degradation. ILLIADS-R combined with iRULE can be used as a reliability estimation system for very large circuits, where iRULE serves as the front-end tool to pinpoint the hot spots and then ILLIADS-R or some other circuit level simulators can be used for more accurate analysis of hot spots and neighboring circuits.

In the next section, the hot-carrier-induced mechanisms leading to oxide damage in MOSFET devices, physical models for elastic scattering and impact- mechanisms and the interface trap generation model will be briefly reviewed. This model forms the basis of the macromodel development and the iRULE reliability evaluation approach.

1.1 Hot-carrier induced device and circuit degradation

The hot-carrier induced degradation of MOS transistors is caused by the injection of high-energy electrons and holes from the channel into the gate oxide region near the drain. The "damage" is in the form of localized oxide charge trapping and/or interface trap generation, which gradually builds up and permanently changes the oxide-interface charge distribution [12], [13], [14]. There are two mechanisms which are mainly responsible for the injection of charged particles into gate oxide. They are (i) elastic scattering of channel hot-electrons, and (ii) creation of high-energy electron-hole pairs by impact ionization [9], [15].

While charge trapping is attributed to electrons injected into the gate oxide through

elastic scattering, the interface trap generation process is primarily due to impact ionization. Recent experimental results show that the oxide damage due to charge trapping occurs mainly under the condition $V_{DS} \approx V_{GS}$ which corresponds to the bias condition for peak electron injection current by elastic scattering, while interface trap generation is observed under the stress bias condition $V_{DS} \approx 2V_{GS}$ where a peak impact ionization current is reached [13], [14]. Thus, it has been suggested that interface traps are predominately created by electrons and holes injected through impact ionization.

When the nMOS transistor is operating in the saturation region, a certain percentage of the hot electrons moving horizontally along the channel create electron-hole pairs by collisions near the drain. The holes created by this process are collected by the substrate, creating the drift component of the substrate current. While most of the created electrons are attracted toward the drain junction, some high-energy electrons and holes are injected into the gate oxide, and contribute to the degradation of gate oxide. Since the majority of holes created by this process constitute the rather easily measurable substrate current I_{SUB} of the MOSFET, the substrate current has been considered a reliable and convenient monitor of the amount of hot-carrier degradation in nMOS transistors [16], [17].

A general expression for the substrate current is found as [9]

$$I_{SUB} = C_1 \, I_{DS} \, \exp\left(-\frac{\Phi_i}{q\lambda E_m}\right) \tag{1}$$

Here Φ_i is the impact ionization energy, Φ_i/qE_m is the distance that an electron must travel in the electric field E_m to gain energy Φ_i, and $\exp(-\Phi_i/q\lambda E_m)$ is the probability of an electron traveling a sufficient distance to gain energy Φ_i or more without collision. Since I_{DS} is the rate of electron flow, $I_{DS}\exp(-\Phi_i/q\lambda E_m)$ is thus the rate of supply of hot electrons possessing energies greater than Φ_i. The parameter C_1 has been experimentally determined as $C_1 \approx 2$ and is actually a weak function of the maximum channel electric field E_m [18].

Equation 1 can also be written as

$$I_{SUB} = \frac{\alpha}{A\beta} E_m I_{DS} \exp\left(-\frac{\beta}{E_m}\right) \tag{2}$$

where α and β are impact ionization coefficients.

As mentioned above, the substrate current I_{SUB} results from the hole generation by the channel hot electrons through impact ionization. The expression for the maximum channel electric field E_m based on a pseudo two-dimensional approach was shown to be [18], [19], [20]

$$E_m = \sqrt{A^2(V_{DS} - V_{DSAT})^2 + E_{sat}^2} \sim A(V_{DS} - V_{DSAT}) \tag{3}$$

where E_{sat} is the critical field for velocity saturation, which is of much smaller interest than that of E_m, and V_{DSAT} is the drain voltage at which the carrier velocity saturates. A is treated as a process-dependent parameter and its value needs to be determined experimentally for each technology. A practical expression for substrate current is described as [21]

$$I_{SUB} = \frac{\alpha}{\beta}(V_{DS} - V_{DSAT}) I_{DS} \exp\left(-\frac{\beta}{A(V_{DS} - V_{DSAT})}\right) \tag{4}$$

1.2 Interface trap generation

It is well-known that the MOSFET degradation is dominated by the generation of acceptor-type interface traps, which are localized in a narrow band near the drain and reduce local mobile carrier density and mobility. One microscopic mechanism is that a hot electron breaks a silicon-hydrogen bond [22]. If the resultant trivalent silicon atom recombines with hydrogen, no interface trap is generated. If the hydrogen atom diffuses away from the interface, a new interface trap is generated. The interface traps responsible for device degradation are generated by hot electrons having energies larger than 3.7 eV, which is about the silicon-hydrogen bond energy plus the barrier energy.

The electro-chemical reaction for this process is described as

$$\equiv Si_3H \overset{hot\ e^-}{\rightleftharpoons} Si^{\bullet} + H_i$$

A hot electron breaks the $\equiv Si_3H$ bond to produce $Si^{\bullet}$, trivalent silicon atom, i.e., interface traps, and H_i, interstitial hydrogen atom.

Approximating the high-energy tail of the electron energy distribution by Maxwell-Boltzmann statistics, the portion of the channel current density which consists of electrons with kinetic energies higher than $\Phi_{it,e}$ can be expressed as the bond-breaking current:

$$I_{BB} = \frac{C_2}{W} I_{DS} \exp\left(-\frac{\Phi_{it,e}}{q\lambda_e E_m}\right) \tag{5}$$

where W is the channel width and I_{DS}/W is proportional to the electron density, and E_m represents the maximum lateral electric field along the channel. The theory of hot-carrier degradation suggests that the bond-breaking current be related to the substrate current and the drain current by the following formula [9]:

$$I_{BB} = \left(\frac{C_2 I_{SUB}^m}{W I_{DS}^{m-1}}\right)^n \tag{6}$$

where n is approximately 0.5, and m is around 3. Although the substrate current I_{SUB} was long regarded as an indicator of hot-carrier damage in nMOS transistors, the quantity I_{SUB}^m/I_{DS}^{m-1} has been shown to be a better measure of the interface trap generation process responsible for degradation [9], [11].

The rate of bond breakage by hot electrons may be expressed as KI_{BB}, where K is proportional to the density of the silicon-hydrogen bonds ($\equiv Si_3H$) at the interface. The rate of $Si^{\bullet}$ and H_i recombination may be expressed as $B_p\, N_{it}\, n_H(0)$, where N_{it} is the interface trap density, $n_H(0)$ is the concentration of H_i at the interface and B_p is a process-dependent constant. Thus, the net rate of interface trap generation is [9]

$$\frac{dN_{it}}{dt} = K\, I_{BB} - B_p\, N_{it}\, n_H(0) \tag{7}$$

Also, the rate of interface trap generation is equal to the rate of interstitial hydrogen diffusing away from the interface, which can be approximated with

$$\frac{dN_{it}}{dt} = D_H \, n_H(0)/X_H \tag{8}$$

Here D_H and X_H are the effective diffusion constant and the effective diffusion length of hydrogen, respectively. By eliminating the $n_H(0)$ term in Eq. (7) and Eq. (8), the following differential equation is obtained:

$$\frac{dN_{it}}{dt}\left(1 + B_p \frac{X_H}{D_H} N_{it}\right) = K \, I_{BB} \tag{9}$$

This is the equation that governs the long-term interface trap generation process by hot-electron injection, under static operating conditions.

Hu et al. have approximated the solution of Eq. (9) by the following simple power-law expression [9]:

$$N_{it}(t) = C \, [t \, I_{BB}]^n \tag{10}$$

Here, the exponent n is in the range of 0.5 to 1. This expression has been widely used for estimating the hot-carrier induced interface trap generation in nMOS transistors.

The dynamic behavior of interface trap generation can be described by the quasi-static model [1], [23], [24]

$$\begin{aligned} N_{it}(t) &= C \left[t \, \frac{1}{T} \int_{t_0}^{t_0+T} I_{BB}(\tau) d\tau \right]^n \\ &= C \, [t < I_{BB} >]^n \end{aligned} \tag{11}$$

where $< I_{BB} >$ is the average bond-breaking current during one cycle period, it provides an accurate measure of the dynamic device hot-carrier degradation.

The enhancement of nMOS device degradation reported for dynamic operating conditions will not be considered in the following analysis. The enhanced dynamic degradation has been linked to hot-hole injection into the gate oxide, a mechanism which is not pronounced in

circuits operating with a power supply voltage $V_{DD} \leq 5$ V, as in most CMOS VLSI circuits. Also the relatively less significant degradation of pMOS transistors is not considered in this report. However, the pMOS degradation issue remains as a future research topic especially for deep submicron CMOS technology.

1.3 Circuit performance degradation

An often-practiced reliability evaluation of CMOS technology with respect to hot-carrier induced degradation is based on dc stress (static stress) experiments of single transistors. A 10% transconductance change is usually set as the lifetime criterion. Single transistor parameters such as linear-mode drain current shifts, transconductance, and threshold voltages are monitored for degradation estimation. As described in the previous section, the hot-carrier degradation can be expressed as [9]

$$degradation = \left[\frac{C_2 I_{SUB}^m}{W I_{DS}^{m-1}} t\right]^n \tag{12}$$

where degradation denotes $\Delta I_{ds}/I_{ds0}$, $\Delta g_m/g_{m0}$, ΔV_{th} described above, W is the device channel width, n is approximately 0.5, and m is around 3.

However, since the actual device operation is dynamic in nature, the operation-related lifetime is of more interest. For example, during the switching of a logic gate, the individual transistors are exposed to a sequence of switching gate and drain voltages. As mentioned in the previous section, we know that the ac-stress lifetime can be determined by the so-called quasi-static model [1], [23], [24]:

$$degradation = (\int \frac{C_2 I_{SUB}^m}{W I_{DS}^{m-1}} dt)^n \tag{13}$$

This quasi-static model has been shown to correctly predict ac degradation in the 1 MHz range for inverter-like waveforms. However, it has been reported that the quasi-static model underestimates ac degradation rates at tens of MHz.

2 Macro-model development for hot-carrier resistant CMOS circuits

A parametric macro-model for CMOS inverter circuit hot-carrier degradation was derived, based on the interface trap generation model described in the previous section. Based on this macro-model, hot-carrier degradation has been studied on other CMOS circuit configurations. In the developed hot-carrier related degradation macro-model, the damage is expressed as functions of the the input signal slope and the ratio of transistor size(W_n) to load capacitance(C_L). The dependence of degradation on these circuit parameters is studied. It is shown that the input signal slope(a) has a much stronger influence on degradation than the ratio of transistor size to load capacitance. Design-for-reliability rules for CMOS circuits can also be developed using this macro-model.

2.1 Measure of hot-carrier damage

The hot-carrier related degradation of the I-V characteristics of nMOS transistors is attributed primarily to the generation of localized oxide-interface traps near the drain when the transistor is operating in the saturation region. Interface traps (states) are generated in nMOS transistors by hot electrons and hot holes, which upon injection into the $Si - SiO_2$ interface break the electron pair bonds. The resulting interface trap may easily acquire an electron when the device is in strong inversion and become negatively charged. Thus, the charge distribution in the oxide interface is largely influenced by the existence of the interface traps. Sah [25] has postulated that the breaking of silicon-hydrogen bonds by the hot carriers is the dominant mechanism of trap generation. Thus the bond-breaking current is a good evaluation of the amount of damage. From the theory of hot-carrier degradation, the

following relation can be achieved:

$$I_{BB} \propto \frac{C_2}{W_n} \frac{I_{SUB}^m}{I_{DS}^{m-1}} \tag{14}$$

Since

$$I_{SUB} = \frac{\alpha}{A\beta} E_m I_{DS} \exp\left(-\frac{\beta}{E_m}\right)$$

the bond-breaking current density can be written as a function of terminal voltages, and device parameters as

$$I_{BB} = \left(\frac{A_i}{B_i}\right)^m \frac{I_{DS}}{W_n} (V_{DS} - V_{DSAT})^m \exp\left(\frac{-mB_i l_c}{V_{DS} - V_{DSAT}}\right) \tag{15}$$

Here, A_i, B_i, m, l_c are all constants, W_n is the device width, and $m \approx 3$. Equation (15) is rewritten as

$$I_{BB} = \frac{K_1}{W_n} I_{DS} (V_{DS} - V_{DSAT})^3 \exp\left(\frac{-K_2}{V_{DS} - V_{DSAT}}\right) \tag{16}$$

where K_1 and K_2 are constants. The macro-model developed in the following section will be based on this equation.

The above equations can evaluate the damage only for static stressing conditions. The bond-breaking currents were assumed to be constant over time. However, for the long-term interface charge accumulation under dynamic operating conditions, Leblebici [26] has proposed a method to use the so-called average bond-breaking current over one cycle period for degradation estimation.

The rate equation for generation of interface states (N_{it}) is

$$\frac{dN_{it}}{dt} [1 + B_p \frac{X_H}{D_H} N_{it}] = K < I_{BB} > \tag{17}$$

By assuming that $< I_{BB} >$ is independent of N_{it}, the exact solution of Eq. (17) is

$$N_{it} + \frac{B_p X_H}{2D_H} N_{it}^2 = K < I_{BB} > t \tag{18}$$

where $< I_{BB} >$ represents the average value of the bond-breaking current over one period, D_H and X_H represent diffusion constants, and K and B_p represent process-dependent coefficients. It is clear from this solution that $< I_{BB} >$ provides an accurate measure for the amount of hot-carrier induced damage N_{it} that each MOS transistor experiences.

2.2 Macro-modeling of CMOS inverter degradation

The average bond-breaking current is shown to be a good measure of hot-carrier degradation. Previous simulation results show that the rising input dominates the hot-carrier degradation compared to the falling input for an inverter circuit. This is illustrated in Fig. 1. Thus, the following analysis assumes that all of the degradation occurs during the rising input.

A simple CMOS inverter is shown in Fig. 2 along with typical input and output waveforms. The bond-breaking current occurs only when the transistor is on and operating in the saturation region, i.e., $I_{BB} \neq 0$ when $V_{GS} > V_T$ and $V_{DS} > V_{DSAT}$.

Assume that a ramp input is applied to the circuit:

$$V_{GS}(t) = a\,t + V_T \tag{19}$$

where a represents the input slope at $V_{GS} = V_T$ and is defined as

$$a \equiv \frac{\Delta V}{\tau_{rise}}$$

here, $\Delta V = V_{in} - V_T$, and the time taken for the ΔV change is τ_{rise} as shown in Fig. 2. When the transistor is in saturation, the following two equations are satisfied:

$$I_{DS} = K_3(V_{GS} - V_T)^2 \tag{20}$$

$$V_{DSAT} = V_{GS} - V_T \tag{21}$$

where, $K_3 = \mu_n C_{OX} W_n / 2L$. By combining these two equations with Eq. (16), the following

expression is obtained:

$$I_{BB} = \frac{K_1}{W_n} K_3 \, (V_{GS} - V_T)^2 \, (V_{DS} - V_{GS} + V_T)^3 \, \exp\left(\frac{-K_2}{V_{DS} - V_{GS} + V_T}\right) \tag{22}$$

With input as a function of time, the output voltage V_{DS} can also be found as a function of time.

From the circuit,

$$I_{DS} = -C_L \frac{dV_{DS}}{dt} \tag{23}$$

$$\frac{dV_{DS}}{dt} = -\frac{K_3}{C_L} \, (V_{GS} - V_T)^2 = -\frac{K_3}{C_L} \, a^2 t^2 \tag{24}$$

By integrating from 0 to time t and knowing $V_{DS}(0) = V_{DD}$, we arrive at

$$V_{DS}(t) = \int_{V_{DS}(0)}^{V_{DS}(t)} dV_{DS} = -\frac{K_3}{C_L} \, a^2 \int_0^t t^2 dt$$

Thus,

$$V_{DS}(t) = V_{DD} - \frac{1}{3} \frac{K_3}{C_L} \, a^2 t^3 \tag{25}$$

Now the bond-breaking current equation can be expressed as a function of time:

$$I_{BB} = \frac{K_1}{W_n} K_3 \, (at)^2 \, (V_{DD} - \frac{1}{3} \frac{K_3}{C_L} \, a^2 t^3 - at)^3 \, \exp\left(\frac{-K_2}{V_{DD} - \frac{1}{3}\frac{K_3}{C_L} a^2 t^3 - a\, t}\right) \tag{26}$$

Figure 3 shows a typical bond-breaking current pulse as a function of time t described by Eq. (26).

The average bond-breaking current, which is found to be an accurate measure of hot-carrier induced damage during dynamic operation, is calculated as

$$< I_{BB} > = \frac{1}{T} \int_0^T I_{BB}(t) dt = \frac{1}{T} \left(\int_0^{t'} I_{BB}(t) dt + \int_{t'}^T I_{BB}(t) dt \right) \tag{27}$$

where T is the period of the input signal and t' is the time point when the transistor gets out of saturation region. The result of this integration gives $< I_{BB} >$ as a function of the input slope a, the device size K_3 and thus W_n, and the load capacitance C_L.

Since $I_{BB} = 0$ from t' to T, the actual integration is carried out from 0 to t', and t' can be found by setting

$$V_{DS}(t') = V_{GS}(t') - V_T \tag{28}$$

By substituting Eq. (19) and Eq. (25) into both sides of Eq. (28), it is found that

$$t'^3 + \frac{3C_L}{K_3 a} t' - \frac{3C_L V_{DD}}{K_3 a^2} = 0 \tag{29}$$

One possible real solution for t' is

$$t' = \sqrt[3]{\frac{-q}{2} + \sqrt{(\frac{q}{2})^2 + (\frac{p}{3})^3}} + \sqrt[3]{\frac{-q}{2} - \sqrt{(\frac{q}{2})^2 + (\frac{p}{3})^3}} \tag{30}$$

where

$$p = \frac{3C_L}{K_3\, a}$$

$$q = -\frac{3C_L V_{DD}}{K_3\, a^2}$$

The average bond-breaking current corresponds to the area under the curve shown in Fig. 3. It can be expressed as a function of two design parameters, a and K_3/C_L. Since $K_3 = \mu_n C_{OX} W_n/2L$, the functional relationship between degradation and important circuit parameters in CMOS inverter can be described by

$$< I_{BB} > = \frac{1}{T} \int_0^T I_{BB}(t)dt \; = f\,(a, W_n/C_L) \tag{31}$$

The result is shown in Fig. 4. The figure shows the normalized degradation of the nMOS transistor in a CMOS inverter circuit as a function of input signal slope a and the W_n/C_L ratio, where a is ranging from 1 V/ns $\sim$ 10 V/ns and W_n/C_L from 40 μm/pF $\sim$ 400 μm/pF. It is seen that the degradation is reduced with increasing input signal slope a. This is because the transistor stays in the saturation region for a shorter time. It is also seen that increasing the W_n/C_L ratio leads to less degradation.

Another interesting result is that the dependence of degradation on input signal slope a is stronger than its dependence on the W_n/C_L ratio. Since a, the input signal slope is determined by the previous stage pull-up transistor, the pMOS transistor size in the previous stage is an important factor. The result also shows that the influence of W_n/C_L upon degradation is stronger with longer input rise time (smaller a) and the effect of W_n/C_L becomes negligible at fast input rise time.

2.3 Design-for-reliability rules for CMOS inverter circuits

The macro-model developed in the previous section can thus be used for deriving the following simple design-for-reliability rules for CMOS inverters: (i) lower power supply voltage leads to less degradation; (ii) faster input signal rise time leads to less degradation; (iii) larger W_n/C_L ratio leads to less degradation. The significance of input rise time was not strongly emphasized previously although it proves to be more important than the W_n/C_L ratio. Thus, the pMOS transistor size in a CMOS inverter of the previous stage is an important design consideration.

2.4 Hot-carrier degradation of CMOS logic gates

The macro-model for CMOS inverter degradation can be extended to CMOS logic gates with some minor modifications. Examples of NAND and NOR gates degradation are discussed in the following. It is shown that in both cases, equivalent inverters can be derived and used. This method also agrees with the results of circuit level simulator iSMILE simulation program.

In CMOS NAND gates, only the uppermost nMOS transistor in the series structure experiences hot-carrier induced degradation during rising input. The degradation is found to be significantly larger for the case in which the input voltage of the uppermost transistor

arrives later than the those of the lower transistors. The worst-case scenario is that all of the lower inputs are in the logic state "one" and the uppermost input switches from "zero" to "one."

Figure 5 shows the equivalent inverter circuit configuration in order to estimate the uppermost transistor degradation. The effect of the lower nMOS transistors in series connection is reflected by a voltage source [27]. The pMOS transistor in the equivalent inverter has the same size as the pMOS transistors in a NAND gate, as does the uppermost nMOS transistor.

The degradation of the uppermost nMOS transistor can be calculated using the expressions derived for a simple CMOS inverter circuit in the previous section using the equivalent circuit. It was shown that the degradation can be estimated with Eq. (16) by setting the source voltage at the source end of the uppermost nMOS transistor to approximately 0.2 V instead of 0 V. This confirms the iSMILE circuit simulation results.

In CMOS NOR gates, where the nMOS transistors are connected in parallel between the output node and ground, the drain-to-source voltage of each nMOS transistor is equal to the output voltage. When all inputs are low initially, the output load capacitance is discharged when one (or more) of the inputs is forced from the low to high voltage level. In this case, only the nMOS transistor with the earliest- arriving rising input voltage will experience significant hot-carrier damage, since the other parallel transistors (receiving later-arriving inputs) will not operate in deep saturation due to the already reduced drain voltage. If more than one transistor is turned on simultaneously, the capacitive discharge current is shared among the conducting transistors, thereby reducing the amount of degradation. As an example, if two input signals arrive simultaneously in a multi-input NOR gate, the degradation of both nMOS transistors is equivalent to the degradation of a single inverter transistor with twice the channel width. This is shown in Fig. 6 where A and B have the same input signals, which arrive earlier than that of C. The expressions derived for the degradation of CMOS

inverters can be used for NOR gate structures by modifying the equivalent channel width W_n.

2.5 Design of CMOS scaled inverter chain buffers for hot-carrier reliability

As a result of improving technology, device dimensions and capacitances on VLSI chips have decreased rapidly. However, the dimensions and capacitances of chip packages and printed wiring connections have remained about the same. Since for MOS circuits the propagation delay is directly proportional to capacitance load, using a normal gate or inverter to drive an off-chip load would result in a long propagation delay. Special output buffer circuits employing a cascade of several progressively larger stages can be used to reduce such delay.

In this section, the overall propagation delay for scaled inverter chain buffers is first discussed. Then, the degradation due to hot-carrier effect is analyzed. In order to consider both propagation delay and circuit degradation, a cost function is defined as the product of the two. An optimum value of the scaling factor F is derived for the tradeoff between speed and degradation. This leads to the a design guideline for scaled inverter chain buffers.

2.5.1 Minimum overall propagation delay

From CMOS circuit transient theory, the propagation delay of a CMOS inverter circuit is given by

$$t_p = \frac{1}{2}\left(t_{PHL} + t_{PLH}\right) \tag{32}$$

where t_{PHL} is defined as the time required for the output voltage to fall from V_{OH} to $V_{50\%} = \frac{1}{2}(V_{OL} + V_{OH})$. Similarly, t_{PLH} is the rise time interval needed to charge output capacitance from V_{OL} to $V_{50\%}$, as shown in Fig. 7. The t_{PHL} and t_{PLH} can be obtained from the following

equations [28]:

$$t_{PHL} = \tau_n \left\{ \frac{2V_{Tn}}{(V_{OH} - V_{Tn})} + \ln \left[\frac{4(V_{OH} - V_{Tn})}{(V_{OH} + V_{OL})} - 1 \right] \right\} \tag{33}$$

where

$$\tau_n = \frac{C_{out}}{K_n (V_{OH} - V_{Tn})} \tag{34}$$

and

$$t_{PLH} = \tau_p \left\{ \frac{2|V_{Tp}|}{(V_{OH} - |V_{Tp}|)} + \ln \left[\frac{4(V_{OH} - |V_{Tp}|)}{(V_{OH} + V_{OL})} - 1 \right] \right\} \tag{35}$$

where

$$\tau_p = \frac{C_{out}}{K_p(V_{OH} - |V_{Tp})|} \tag{36}$$

If the CMOS inverter is designed in a completely complementary manner with $K_n = K_p$ and $V_{Tn} = |V_{Tp}| = V_T$, then the propagation delay time is

$$t_p = t_{PHL} = t_{PLH}$$

It is seen that the delay is directly proportional to the load capacitance.

The total propagation delay can be reduced by replacing an inverter with a scaled inverter chain buffer. Figure 8 shows a scaled CMOS inverter chain buffer. The calculation of the optimum scaling factor F for minimum overall delay will be first reviewed for self-sufficiency. Suppose that the input capacitance for this inverter chain is C_G and the output load capacitance is C_L. The number of stages in this chain is N, and F is the *fan-out factor*, the capacitance seen by each stage relative to that seen by the preceding stage. Either N or F is the independent design parameter. Define the propagation delay of one minimum size stage driving another identical stage (F=1) as t_{p0}. Then, the overall delay for the buffer is [29]

$$t_{total} = N F t_{p0} \tag{37}$$

The relation between N, F, and the overall capacitance ratio is given by

$$F^N = \frac{C_L}{C_G} \tag{38}$$

i.e., N must satisfy

$$N = \frac{\ln C_L/C_G}{\ln F} \tag{39}$$

The overall delay is expressed by scaling factor F as

$$t_{total} = \frac{\ln C_L/C_G}{\ln F} F t_{p0} \tag{40}$$

To find the minimum value of total propagation, the first derivative of Eq. (40) with respect to F is set equal to zero.

$$\frac{\partial t_{total}}{\partial F} = 0$$

$$\frac{1}{\ln F} - (\frac{1}{\ln F})^2 = 0$$

$$F = e = 2.72 \tag{41}$$

Thus, at the optimum, the total delay in the output buffer is

$$t_{total} = 2.72 \ln C_L/C_G \, t_{p0} \tag{42}$$

This improves the total propagation delay compared to the single inverter case. The device sizes for the first stage are those typical of the on-chip logic. Each following stage in the output buffer has device channel width F times wider than that of the preceding stage.

2.5.2 Degradation as a function of scaling factor F

The hot-carrier induced degradation for CMOS scaled inverter chain buffers can also be expressed as a function of circuit parameters such as input rising time, transistor size, and load capacitance. But because of the special structure of the buffer, some simplifications can be made, and a direct relation between the average bond-breaking current and the scaling factor F is found.

The input signal slope is defined as

$$a \equiv \frac{\Delta V}{\tau_{rise}}$$

Since t_{p0} represents the propagation delay when the transistor sizes are the same for each stage, i.e., $F = 1$, then $F\,t_{p0}$ is the delay when the scaling factor is F. Thus,

$$a = \frac{\Delta V}{F\,t_{p0}} \tag{43}$$

where ΔV is the voltage change during the time period $F\,t_{p0}$. The W_n/C_L ratio in the macro-model can also be found as a function of F. Figure 9 shows the ith stage in the inverter chain. It is clear that the next stage capacitance is dominated by the gate capacitance, i.e.,

$$C_L = C_{i+1} = C_{OX}\,(W_p + W_n)_{i+1}\,L \tag{44}$$

where C_{OX} is the oxide capacitance. Assuming equal channel length L for all transistors, if the channel widths for pMOS and nMOS transistors on stage i are $W_{p,i}$ and $W_{n,i}$, respectively, then for this scaled inverter chain, the next stage transistor widths are $W_{p,i+1} = F\,W_{p,i}$ and $W_{n,i+1} = F\,W_{n,i}$ where F is the scaling factor. Thus,

$$\frac{W_n}{C_L} = \frac{W_{n,i}}{C_{OX}\,(W_p + W_n)_{i+1}\,L} \tag{45}$$

For a completely complementary CMOS circuit with $K_n = K_p$, we have

$$\mu_n\,W_n = \mu_p\,W_p \tag{46}$$

Equation (45) can be replaced as

$$\begin{aligned} \frac{W_n}{C_L} &= \frac{W_{n,i}}{W_{n,i+1}}\,\frac{1}{(1 + W_{p,i+1}/W_{n,i+1})\,C_{OX}\,L} \\ &= \frac{W_{n,i}}{W_{n,i+1}}\,\frac{1}{(1 + \mu_n/\mu_p)\,C_{OX}\,L} \end{aligned} \tag{47}$$

It is seen that

$$F = \frac{C_{i+1}}{C_i} = \frac{W_{i+1}}{W_i} = \frac{W_{n,i+1}}{W_{n,i}} = \frac{W_{p,i+1}}{W_{p,i}}$$

Thus, the ratio of W_n/C_L can finally be expressed as a function of F and some constants:

$$\frac{W_n}{C_L} = \frac{1}{F}\,\frac{\mu_p}{(\mu_p + \mu_n)\,C_{OX}\,L} \tag{48}$$

Applying Eq. (43) and Eq. (48) to the macro-model Eq. (31). the relationship between degradation and scaling factor F is obtained. Figure 10 shows that degradation is linearly dependent on F. This result can be applied to any stage, and it is clear that $< I_{BB} >$ depends only on the scaling factor F. Once F is chosen, all stages experience the same amount of degradation.

2.5.3 Design tradeoff between speed and degradation

An important design goal for the scaled inverter chain buffer would be to reduce total propagation delay and hot-carrier degradation at the same time. From the analysis of previous sections, two important results are found out. First, each single inverter within the inverter chain buffer experiences the same amount of hot-carrier degradation, which depends only on the scaling factor F. Second, the degradation is a linear function of F. Let us define a cost function to be the product of the total propagation delay and the inverter degradation, i.e., let the product

$$P = < I_{BBi} > t_{total} \tag{49}$$

be a cost function, where $< I_{BBi} >$ represents the degradation on the ith stage and t_{total} represents the total propagation delay for the inverter chain. In this way, the minimization of the product tends to provide a good compromise between the degradation of the buffer and the total switching delay. Suppose the amount of degradation on one stage equals $< I_{BB0} >$ when F=1, i.e., an inverter chain with all of the inverters having the same size. Then, for an arbitrarily chosen F, the degradation on this stage will be

$$< I_{BBi} > = F < I_{BB0} > \tag{50}$$

Using Eq. (37),

$$t_{total} = N F t_{p0} \tag{51}$$

the cost function becomes

$$P = F < I_{BB0} > N\ F\ t_p = N\ F^2 < I_{BB0} > t_p \tag{52}$$

Equation (39) is shown again here

$$N = \frac{\ln C_L/C_G}{\ln F} \tag{53}$$

The minimum of Eq. (52) subject to the constraint in Eq. (53) is found by taking the partial derivative of P with respect to F, i.e..

$$\frac{\partial P}{\partial F} = 0 \tag{54}$$

$$\begin{aligned} \frac{\partial P}{\partial F} &= \frac{\partial}{\partial F}\left(\frac{F^2}{\ln F}\ln\frac{C_L}{C_G} < I_{BB0} > t_{p0}\right) \\ &= \frac{(2F\ln F - F)}{(\ln F)^2}\ln\frac{C_L}{C_G} < I_{BB0} > t_{p0} = 0 \end{aligned} \tag{55}$$

The solution for F is

$$F = \sqrt{e} = 1.65 \tag{56}$$

The result indicates that the degradation-delay product is minimized when the *scaling factor* F equals the square root of the base of natural logarithms. The linear dependence of degradation on F is shown in Fig. 11 along with the normalized propagation delay and degradation product cost function. Notice that the propagation delay curve for a typical scaled inverter chain reaches its minimum if the scaling factor $F = e$. The dotted line is the normalized cost function and its optimal value is reached at $F = \sqrt{e} = 1.65$.

In general, two different scaling factor values may satisfy a given maximum delay requirement. In such a case, Fig. 11 indicates that an inverter chain with a smaller scaling factor, i.e., with a larger number of stages, is expected to be more reliable than an inverter chain with a larger scaling factor and with the same overall propagation delay.

The figure also shows that although $F = 1.65$ is the optimum value for the minimization of the delay-degradation product, the propagation delay is very high. On the other hand, the propagation delay is not changing much beyond the point $F = e$, even at about $F = 5$. Based on this observation, using a larger scaling factor is not advisable, since the degradation is linearly increasing with F. In view of both the delay and the hot-carrier degradation, $F = 2$ may be a good choice.

2.6 Design of nMOS transmission gates for hot-carrier reliability

The transmission gate is often used in logic circuits. It is found that in nMOS transmission gate circuits, the transistor experiences significantly more hot-carrier damage during the charge phase than during the discharge phase [30]. As the transmission gates are turned on and off, charge sharing occurs between the source and drain nodes. The worst-case degradation is considered when the drain is already in logic "one," i.e., at V_{DD}, and the load capacitance is connected to source. The source voltage is found when a ramp input is applied to the gate, and the macro-model is used for degradation estimation.

Figure 12 shows an nMOS transmission gate circuit with the drain end is at V_{DD} and a source voltage equals to zero. A ramp input is applied to the gate with slope a during the charge-up phase.

$$V_G(t) = at + V_T \tag{57}$$

thus,

$$V_{GS} - V_T = V_G - V_S - V_T = a\,t - V_S \tag{58}$$

As the transistor starts to turn on, the load capacitance is being charged up.

$$\frac{dV_S}{dt} = \frac{K_3}{C_L}(V_{GS} - V_T)^2 = \frac{K_3}{C_L}(a\,t - V_S)^2 \tag{59}$$

The above equation can be rewritten as

$$\frac{dV_S}{dt} = \frac{K_3}{C_L}V_S^2 - \frac{2\,K_3}{C_L}a\,t\,V_S + \frac{K_3}{C_L}a^2\,t^2 \tag{60}$$

which is a Ricatti differential equation. The solution for this equation is found out to be:

$$V_S = \sqrt{\frac{aC_L}{K_3}}\,\frac{\exp(-2\,t\sqrt{\frac{aK_3}{C_L}}) - 1}{\exp(-2\,t\sqrt{\frac{aK_3}{C_L}}) + 1} + a\,t \tag{61}$$

The hot-carrier degradation of a transmission gate can be measured using the average bond-breaking current. Equation (16) is written again here:

$$I_{BB} = \frac{K_1}{W_n}I_{DS}(V_{DS} - V_{DSAT})^3\exp(\frac{-K_2}{V_{DS} - V_{DSAT}}) \tag{62}$$

During charging of the load capacitance, the drain current of the nMOS transmission gate can be expressed as

$$\begin{aligned} I_{DS} &= K_3(V_{GS} - V_T)^2 \\ &= K_3(a\,t - V_S)^2 \\ &= K_3\left(\sqrt{\frac{aC_L}{K_3}}\frac{exp(-2\,t\sqrt{\frac{aK_3}{C_L}}) - 1}{exp(-2\,t\sqrt{\frac{aK_3}{C_L}}) + 1}\right)^2 \end{aligned} \tag{63}$$

and "the depth of saturation" can be found as a function of t:

$$V_{DS} - V_{DSAT} = V_D - V_S - (V_{GS} - V_T) = V_{DD} - a\,t \tag{64}$$

Thus, the bond-breaking current is expressed only as a function of the input signal slope a, the transistor size, the load capacitance and time t, which will be eliminated after calculating the average value of bond-breaking current:

$$I_{BB} = K_3\,\frac{K_1}{W_n}\,(V_{DD} - a\,t)^3\,\exp\left(\frac{-K_2}{V_{DD} - a\,t}\right)\cdot\left(\sqrt{\frac{aC_L}{K_3}}\,\frac{\exp(-2\,t\sqrt{\frac{aK_3}{C_L}}) - 1}{\exp(-2\,t\sqrt{\frac{aK_3}{C_L}}) + 1}\right)^2 \tag{65}$$

where

$$K_3 = \frac{1}{2}\mu_n C_{OX} \frac{W_n}{C_L}$$

and

$$K_3 = K_3' W_n$$

With K_3' being a constant independent of transistor channel width, Eq. (65) now becomes

$$I_{BB} = K_3' K_1 (V_{DD} - a\,t)^3 \exp\left(\frac{-K_2}{V_{DD} - a\,t}\right) \cdot \left(\sqrt{\frac{a\ C_L}{K_3' W_n}} \frac{\exp(-2\,t\sqrt{a\,K_3' \frac{W_n}{C_L}}) - 1}{\exp(-2\,t\sqrt{a\,K_3' \frac{W_n}{C_L}}) + 1}\right)^2 \quad (66)$$

The average value is found the same way as for the CMOS inverter case:

$$< I_{BB} >= \frac{1}{T}\int_0^T I_{BB}(t)dt \quad (67)$$

Hence, the resulting hot-carrier induced damage during the charge-up phase is described by a parametric closed-form expression, which is a function of the input signal slope a and the W_n/C_L ratio.

$$< I_{BB} >= f(a, W_n/C_L) \quad (68)$$

Figure 13 shows the expected degradation of the nMOS transmission gate transistor as a function of a and W_n/C_L. In this case, the input signal slope is changing from 1.0 V/ns to 10.0 V/ns, and W_n/C_L ratio from 40.0 μm/pF to 400.0 μm/pF.

2.7 Summary

This section has presented a parametric macro-model approach for the evaluation of hot-carrier-related degradation of MOS transistors operating in digital logic circuits. Previously, the simulation approach has been used for reliability analysis, but it is inefficient for reliability assessment of very large scale integrated circuits. Geometry-based macro-models for reliability estimation have been developed. Design parameters that are critical to circuit reliability can be identified from the circuit layout and can be used as parameters of macro-models.

Some design-for-reliability rules have been developed, based on the new macro-models for early diagnosis of potential reliability problems in CMOS circuits.

The macro-models for MOS circuits degradation are based on the well-known interface trap generation model for hot-carrier degradation. It is found out that the average bond-breaking current $< I_{BB} >$, which is an accurate measure of the amount of hot-carrier degradation when the transistor is operating in saturation region, can be expressed as a function of layout geometry related parameters such as the nMOS transistor channel width W_n, the load capacitance C_L, and the input signal rise time a.

For CMOS inverter circuits driven by a ramp input, hot-carrier induced damage occurs primarily during the rising input phase. The degradation is a function of two designable circuit parameters, the input signal slope a, and the ratio of the nMOS transistor channel width to load capacitance (W_n/C_L). In fact, the input slope a is primarily determined by the (W_p/C_L) ratio of the previous inverter stage. It is seen that $< I_{BB} >$ increases with decreasing a and with decreasing (W_n/C_L). Another important result is that the influence of the input signal slope a upon degradation is stronger than that of the (W_n/C_L) ratio.

The developed macro-models can thus be used for deriving the following simple design-for-reliability rules for CMOS inverters:

(i) Smaller input signal rise time leads to less degradation: $< I_{BB} > \propto 1/a^{0.8}$;

(ii) Larger (W_n/C_L) ratio leads to less degradation: $< I_{BB} > \propto 1/(W_n/C_L)^{0.3}$.

The pMOS transistor size of the *previous* stage must also be regarded as an important design consideration for long-term reliability, because of the strong influence of input rise time upon dynamic degradation.

In scaled CMOS inverter chain buffer circuits, the nMOS and pMOS transistor channel width to channel length ratios W/L of each inverter stage are F times larger than that of the proceeding stage. It is found that each nMOS transistor in the chain experiences the

same amount of hot-carrier degradation. This amount of degradation is a linear function of the scaling factor F. Considering the tradeoff between the speed and the performance degradation during switching, a cost function is defined as the product of the hot-carrier degradation and the total propagation delay. An optimum value is obtained for this cost function when F equals $\sqrt{e}$.

In CMOS NOR gates, only the nMOS transistor with the earliest arriving rising input will experience significant hot-carrier related degradation. The macro-model for CMOS inverters can be used for NOR gate structures by modifying the equivalent channel width W_n. In CMOS NAND gates, only the uppermost nMOS transistor in the series structure experiences hot-carrier induced degradation during rising input. The degradation is found to be significantly larger for the case in which the uppermost input arrives later than the lower inputs.

The nMOS transmission gate circuit is also studied. The hot-carrier damage appears much more extensive during the *charge up* phase than during the *discharge* phase. The degradation can also be expressed as a function of a, the input signal slope and the (W_n/C_L) ratio.

One application of the derived design rules for hot-carrier degradation is the development of a rule-based reliability diagnosis tool. The tool should be able to determine how the overall circuit performance is affected as a result of component aging; it should also be able to identify which circuit components are mainly responsible for the degradation of circuit performance. This tool can basically be a rule-checking program. It reads in a circuit description file, such as a SPICE type input file, and gives the expected hot-carrier degradation of critical transistors as output.

The design of a tester chip is necessary for on-chip verification of the design-for-reliability rules for hot-carrier degradation. The testing results will give feedback and calibration for

the developed macro-model.

The next two sections will discuss these issues.

3 Tester chip design

A tester chip has been designed for hot-carrier reliability testing. The chip has been fabricated by MOSIS. The technology is CMOS n-well 1.2μm LDD structure. The test structure contained four different parts in order to measure device and circuit hot-carrier degradation for CMOS circuit and for nMOS passgates.

According to the developed macromodel, the hot-carrier degradation of CMOS inverter and nMOS passgate is a function of the signal input slope, the transistor sizing, and the capacitive loading. The tester chip is so designed to verify the macormodel, and the testing results may further provide information on design guidelines and suggest design rules for hot-carrier reliability consideration.

The tester chip is made up of four separate test structures, each with 24 accessing pads.

- The first structure contains seven single transistors, with four nMOS and three pMOS transistors. Each transistor can be accessed by its source, drain, gate, and substrate node shown in Fig. 14.

- The second test circuit is made up of CMOS inverters with different transistor sizing and CMOS passgates. The purpose of this design is to find out how degradation depends on W/C_L ratio while having the same input slope. Notice there are three CMOS passgates attached to each inverter stage. They serve different purposes. When stress is applied to the inverters, the input passgates are on while the output passgates are off in order to avoid any noise from the output pad. When testing is performed, the input passgates are off and the output passgates are on, so is the passgate to access the

gate node of CMOS inverter. In this way all the nodes of the transistor under stress can be accessed. The first stage inverter is used to generate the input slope. Figure 15 is the structure of the circuit.

- The third circuit consists of three similar structures, one inverter stage to generate input slope for the next stage under testing. Each testing stage receives different rising slope because of the capacitance added to decrease the slope and to eliminate the accessing passgate to increase the slope. SPICE simulations show that the slope can not be made very steep because of the parasitics. This is shown in Fig. 16.

- The fourth chip Fig. 17 is for nMOS passgate testing with same input slope and different W/C_L ratios. Also in this chip are EOS/ESD test circuits with I/O protection.

Another design is a test circuit containing passgates and an XOR gate shown in Figure 18. This design was fabricated by AT&T Bell Labs. Initial measurements were performed. The purpose of this design is to find out how the effect of the hot-carrier degradation on device will affect circuit performance. The idea is to detect any signature in the output of a circuit due to hot-carrier induced degradation of a transistor. The designed circuit consists of two nMOS passgates each driven by an inverter stage. One passgate is to be stressed while the other one is kept fresh. During the measurement, the output of the two passgates are fed into an XOR gate. If both passgates are fresh, they should have the same delay properties and thus the output of the XOR will always be zero. But if a passgate is aged, then the delay through it will be different from the delay through the fresh transistor. Thus, the output of XOR will become one. Through iSMILE simulation and SPICE simulation, we found a spike at the output of the XOR, and its height is a function of load capacitance of the XOR gate, the pull up ability of the XOR, which is the ratio of W_p/W_n and the delay difference of the two signal paths.

The circuit is implemented using an AT&T standard cell library, with proper transistor sizing.

4 Reliability rule checking program–iRULE

A new macromodel-based reliability diagnosis program (iRULE) is presented for design of hot-carrier resistant VLSI circuits. iRULE detects first the potentially critical transistors by using netlist- and geometry-based rules and then determines the criticality of hot-carrier degradation in those transistors by using macromodels. The algorithmic complexity of iRULE is shown to be linear with respect to the number of logic gates. iRULE is capable of performing reliability diagnosis for standard CMOS combinatorial and sequential circuits including transmission gates. However, it can not handle CMOS circuits with nMOS loading or resistive loading at this stage.

The current version of iRULE provides a circuit-level, worst-case analysis of hot-carrier related reliability problems. However, the rule-based reliability diagnosis framework is extendable to encompass other sources of reliability hazards such as electromigration and electrical overstress, as illustrated in Fig. 19. The significant benefits that iRULE offers, over the traditional simulation-based reliability diagnosis approaches, are:

1. Ability to analyze very large circuits with more than one million transistors on a workstation,

2. Very short execution time, and the algorithmic complexity of iRULE is *linear* in the number of transistors in the circuit.

These benefits can be exploited by adopting iRULE as the front-end tool in a hierarchical design-for-reliability system. The transistor-level circuit description (for example, the output

of a layout extraction tool) is first given to iRULE, and iRULE quickly highlights the critical transistors in the design. The designer can then take this information to:

1. Carry out more detailed analysis of the relevant (small) parts with greater accuracy using switch-level reliability simulation tools such as ILLIADS-R [5] or circuit-level reliability simulators such as iSMILE [2, 3] and BERT [6], in order to fine tune the circuit in accordance with the desired reliability specifications, and/or

2. Re-design the relevant parts of the circuit.

The macro-models developed in [7, 8] form the basis for iRULE, being responsible for some of the netlist- and geometry-based rules employed by iRULE in determining the potentially critical transistors, and for the look-up table of hot-carrier degradation values—given as a function of certain circuit parameters.

The accuracy of these macro-models have been verified by comparing the predicted amount of degradation in the nMOS transistor (using the macro-models) with circuit-level reliability simulation results using iSMILE (Fig. 20) [2, 3]. The observed deviation of about 10% between the macro-model results and the circuit simulation results was found to be due to the numerical inaccuracies introduced during the integration in the evaluation of the macro-models. These simple macro-models accurately represent the parametric dependency of device degradation upon the rising input slope and the W_n/C_L ratio.

4.1 The iRULE program

The procedural description of the program iRULE is given in Fig. 21. The algorithmic complexity of iRULE is linear in the number of logic gates, as seen from the description in Fig. 21. Two input files are required: the circuit description file and the device degradation look-up table file. The format of the circuit description file is the same as that of SPICE with

the following exceptions: (i) Specification of input waveforms is not required, as iRULE is input-vector independent. (ii) The input file does not need .TRAN and associated command lines, as iRULE does not run transient circuit simulations. The drain areas and drain perimeters of the transistors are used to determine the parasitic drain capacitances, which would be available (via an extraction program) from the layout provided. Alternatively, the user can provide numerical estimates of these parameters. The ability to utilize circuit extraction results is an attractive feature of iRULE for reliability estimation, which is derived from the simplicity of the geometry-based macro-models used in iRULE. The output of the circuit extraction tool, when the extractor is run on the given design layout, has virtually all the information required by iRULE. As is the case with most parameters in iRULE, default values are used unless explicitly defined. The device degradation look-up table, which must be provided as an input file, is the table of $\langle I_{BB} \rangle$ values as a function of rising input signal slope a and ratio W_n/C_L. Given a circuit description file, the W_n/C_L of each nMOS transistor in a CMOS inverter can be calculated and so does the input signal slope a. The look-up table is generated by varying a and W_n/C_L in a certain range and calculating $\langle I_{BB} \rangle$ for each point of a and W_n/C_L according to Eq. (31). Once the a and W_n/C_L are found out for a certain transistor, the program goes into the table and finds the $\langle I_{BB} \rangle$ value for this transistor by interpolation or extrapolation. Note, a separate table is needed for pass-gates. The output file gives the $\langle I_{BB} \rangle$ value, i.e., the amount of expected hot-carrier-induced stress, for each critical transistor, and the mean $\langle I_{BB} \rangle$ averaged over all critical transistors.

iRULE first reads in the circuit description file and creates an internal representation of the given circuit. iRULE determines the worst-case equivalent W/L ratio for each pMOS block. This parameter is needed to compute the slope a of the rising input into gates being driven by the logic gate under consideration. The effective load capacitance C_L as seen by each gate output is computed next. C_L is needed to calculate the W_n/C_L ratio, which

serves as one of the two indices into the look-up table, and to compute the rising output slope of each logic gate. Once the ratio W_n/C_L and the input rising slope a (which are used to determine the degradation through interpolation via the look-up table) have been computed, iRULE traverses through the internal circuit representation to determine the critical transistors using the look-up table. Accurate interpolation is done using a technique employing 3-dimensional isoparametric shape functions [31].

The computational cost of iRULE increases *linearly* with the number of transistors in the given circuit, unlike simulation-based tools which have a non-linear cost-dependence on circuit size. This allows iRULE to run on large circuits in a reasonable amount of time. The results are shown in the next section.

4.2 Validation and time complexity of iRULE

To validate iRULE, its $\langle I_{BB} \rangle$ results have been compared with the $\langle I_{BB} \rangle$ values obtained from running the circuit-level simulator iSMILE [2, 3] on test circuits with varying input slopes, transistor sizes and capacitive loadings. Table 1 is a compilation of some of our comparison results. Both circuits presented are 5-stage inverter chains. The one large (negative) percentage difference, in Table 1, is due to the fact that the rising signal slope a at the gate output is calculated using a closed-form expression (as given above) which does not accurately take into account the contribution of previous stages to the rise-time of the input signal. However this problem can be corrected by empirically modifying the above-mentioned formula.

The results in Table 2 demonstrate that iRULE is very fast, even for the circuit consisting of more than one million transistors. These results are for iRULE running on a SUN SPARCstation 2, with 64 MByte physical memory and 256 MByte virtual memory. The circuit "eld" is an industry circuit, and the other circuits—with the prefix 'R'—are circuits

Ckt	Stage	Rise-time (nS)	W/C ratio (m/F)	iRULE	iSMILE	% Difference
inv5	stage1	1.0	1.976e+08	1.002e-08	0.858E-08	16.8
	stage2		1.976e+08	6.869e-09	0.686E-08	0.1
	stage3		1.976e+08	6.869e-09	0.652E-08	5.4
	stage4		1.976e+08	6.869e-09	0.667E-08	3.0
	stage5		4.545e+07	7.192e-09	0.702E-08	2.4
inv5	stage1	1.5	1.976e+08	1.470e-08	0.125E-07	17.6
	stage2		1.976e+08	6.869e-09	0.850E-08	-19.2
	stage3		1.976e+08	6.869e-09	0.664E-08	3.5
	stage4		1.976e+08	6.869e-09	0.667E-08	3.0
	stage5		4.545e+07	7.192e-09	0.702E-08	2.4

Table 1: iRULE compared with iSMILE.

built by adding replications of the "eld" circuit. Fig. 22 is a plot of the data in Table 2.

The final column in Table 2 gives the ratio of CPU milliseconds per transistor. From this column we see that by the time a quarter-million transistors is reached, the limit on the workstation's available physical memory is beginning to become a constraint—in that swapping between physical memory and disk is starting to become more frequent. In addition, the C memory allocation routine `malloc()` starts to incur higher overhead. This is the reason for the slight non-linear trend for those circuits containing over a quarter-million transistors. From the final column in Table 2 one can observe that for the 1.1 million transistor circuit, the use of virtual memory capacity is pushed to its limit. We observed that the memory allocation routine `malloc()` began to fail when iRULE was run on circuits larger than about 1.1 million transistors. Newer workstations, for example, a SPARCstation 10 with the capacity for a half-GByte (512 MByte) of physical memory, would be able to *efficiently* swap even when using up to 1.25 GByte virtual memory. Therefore, it is clear that iRULE, running on a more powerful workstation, should be able to efficiently handle circuits with even over 10 million transistors. Table 2 shows what we expect in theory: iR-

Circuit name	Number of				CPU secs.	Ratio: mSecs/Transistor
	transistors	gates	inputs	outputs		
eld	227	49	22	9	0.5	1.98
R1	454	98	35	9	0.9	1.87
R2	908	196	61	9	1.5	1.69
R3	1816	392	113	9	2.9	1.61
R4	3632	784	217	9	5.7	1.58
R5	7264	1568	425	9	11.2	1.54
R6	14528	3136	841	9	22.4	1.54
R7	29056	6272	1673	9	44.9	1.54
R8	58112	12544	3337	9	89.8	1.55
R9	116224	25088	6665	9	182.5	1.57
R10	232448	50176	13321	9	389.3	1.67
R11	464896	100352	26633	9	878.0	1.89
R11.x	619710	133770	35499	9	1255	2.03
R11.y	805396	173852	46133	9	1723	2.14
R12	929792	200704	53257	9	2087	2.24
R12.x	1100042	237454	63007	9	2600	2.36

Table 2: iRULE run-times, on a SPARCstation 2.

ULE's run-time performance is linear with respect to the size of the circuit, unless memory constraints are encountered. The linear run-time complexity is to be expected, as iRULE has linear algorithmic complexity, as shown above.

Since iRULE does not use simulation, it is very fast, as demonstrated in Table 2 and Fig. 22. For comparison, the fast simulation-based reliability tool ILLIADS-R [5] took 240 seconds on a 3,786–transistor circuit on a SPARCstation, whereas iRULE ran in just 11.2 seconds for a 7,264–transistor circuit. These circuits cannot, however, be simulated by circuit-level reliability simulators.

4.3 iRULE application—hot-carrier circuit A.C. lifetime evaluation

iRULE has been successfully technology-transferred to AT&T Bell Labs at Allentown, PA. The degradation macromodels based on the well-known substrate current model for hot-carrier degradation were implemented on the AT&T IC design platform. iRULE predicted circuit lifetimes based on the extrapolated transistor DC lifetimes(10% g_m degradation) and the circuit duty cycle considering the worst case scenario. It reads SPICE-like input translated from ADVICE(AT&T simulation tool) circuit description file. Lifetime predictions were made for a variety of circuits such as ALU, multipliers, comparators and other circuits constructed from 0.9μm and 0.5μm CMOS standard cell libraries. The predicted circuit AC lifetime for a 0.5μm 16x16 multiplier was longer than 30 years under operation at 100MHz with 5.5V power supply. The 1042BG circuit, a transmitter receiver circuit with more than 90,000 transistors and extracted from the layout, was also successfully analyzed.

5 Fast dynamic reliability simulation

Various approaches have been made previously for modeling the hot-carrier related degradation mechanisms in MOS transistors, and for estimating their influence upon long-term circuit reliability. Hu *et al.* have proposed a semi-empirical relationship between the interface trap density and the substrate current under static operating conditions [9]. Later, the same approach has been adapted by Kuo *et al.* to circuit-level reliability simulation, for estimation of device lifetime under quasi-static operating conditions [1]. The hot-carrier reliability simulator HOTRON developed by Aur *et al.* also follows a similar approach [16]. Quasi-static lifetime estimation is based on the assumption that the stress conditions of each transistor remain unchanged during dynamic operation. The authors have proposed a

reliability simulation approach which accounts for the gradual variation of dynamic stress conditions [32, 3]. A repetitive simulation scheme has been adopted to ensure accurate prediction of circuit-level degradation. However, the detailed circuit simulation, needed for determining the stress conditions of individual devices, restricts the computational efficiency of reliability simulation approaches for very-large scale integrated circuits.

The reliability simulation tool presented here incorporates a new second-order polynomial equation model for hot-carrier damaged MOS transistors for computational efficiency. A repetitive simulation scheme proposed earlier by the authors [3] has been adopted to ensure accurate prediction of the circuit-level degradation process under dynamic operating conditions. Unlike many other fast timing simulators, ILLIADS is capable of providing analogous voltage waveforms associated with each circuit primitive. A typical ILLIADS simulation strategy of two cascaded gates is shown in Fig. 23. This unique feature enables fast and accurate calculation of hot-carrier damages without significant increase of simulation time. Simulation results demonstrated that the proposed reliability simulation using ILLIADS can provide a computational speed-up of several orders of magnitude, while achieving an accuracy comparable to that of transistor-level simulators. Thus, ILLIADS simulation tool offers the capability for simulation of hot-carrier damages in very-large scale MOS circuits. The fast simulation capability also enables the implementation of statistical reliability prediction techniques, which did not exist in previous hot-carrier simulation tools [1, 16, 3].

5.1 Models for hot-carrier induced damage

The framework for a fast hot-carrier reliability timing simulation tool involves (i) modeling the gate oxide degradation in the MOS transistor as a function of its operating conditions, (ii) modeling the behavior of the MOS transistor with localized oxide damage, (iii) simulation of gate oxide degradation during long-term circuit operation, and (iv) determination of the

overall circuit performance after hot-carrier stress. The integration of these components into a simulation framework is the prerequisite for accurate prediction of the long-term circuit reliability characteristics.

The hot-carrier related degradation of the I–V characteristics of MOS transistors is attributed primarily to the generation of localized oxide-interface traps near the drain [14]. The trap generation process can be characterized by the following simplified bond-breaking current model.

$$I_{BB} = \left[\frac{A_i}{B_i}\right] I_{DS} \cdot (V_{DS} - V_{SAT})^{\alpha} \tag{69}$$

with $\alpha \approx 2.3$.

The interface trap generation model given by Eqs. (69) and (17) can be applied for predicting the long-term dynamics of hot-carrier induced device degradation as long as the associated stress conditions, i.e. the terminal voltage waveforms remain unchanged. Therefore, the use of the above degradation models must be restricted to shorter time intervals in which the change of stress conditions may be neglected. To extend the simulation of device degradation mechanisms over longer time intervals, the stress conditions must be updated periodically by simulating the circuit operation with damaged nMOS transistors. This task emphasizes the need for a simple and accurate device model to represent the behavior of the transistor with hot-carrier induced oxide damage.

To enable the accurate simulation of circuit behavior during and after the process of hot-carrier induced degradation, a simple empirical damaged MOSFET model has been newly developed. Given the amount of hot-carrier induced damage by a few parameters such as N_{it}, and the terminal voltages, this model is capable of accurately representing the current-voltage characteristics of the locally damaged nMOS transistor. The new model equations

for the linear and saturated operating regions are given below.

$$I_{DS} = \begin{cases} \frac{K_p W}{2L}\left(2A(V_{GS}-V_T)V_{DS} - BV_{DS}^2\right) & \text{for } V_{DS} < V_{DSAT} \\ \frac{K_p W}{2L}\left(C(V_{GS}-V_T)^2 + \lambda V_{DS}(V_{GS}-V_T)\right) & \text{for } V_{DS} \geq V_{DSAT} \end{cases} \tag{70}$$

Here, the saturation voltage V_{DSAT} is defined as

$$V_{DSAT} = \gamma \frac{A}{B}(V_{GS} - V_T) \tag{71}$$

Note that the current continuity condition at the linear-saturation region boundary dictates the following relationship between the model parameters.

$$C = \gamma \frac{A}{B}(2A - \gamma A - \lambda) \tag{72}$$

It can be seen that the proposed simple transistor model depends on six independent modeling parameters, i.e. A, B, K_p, V_T, λ and γ. The current equations are second-order polynomials of the terminal voltage variables V_{GS} and V_{DS} in both the linear region and the saturation region, and therefore can be handled by ILLIADS efficiently. The model parameters can be updated during the reliability simulation as a function of the hot-carrier damage level (N_{it}) to simulate the I–V characteristics of the damaged MOS transistor either by using table lookup or empirical formula.

In previous reliability simulation approaches, complicated device modeling equations and extensive updating of device parameters have been used. The new capability to simulate damaged transistors without extensive parameter extraction before and after hot-carrier stress is also valuable for the overall simplicity of the proposed simulation tool.

Figure 24 shows the I–V curves generated by using this simple model at different stress levels.

5.2 Reliability simulation algorithm

The fast dynamic reliability simulation is similar to the timing simulation except that only the transistor model parameters become different after aging.

The ILLIADS program first partitions a circuit into DCCBs (DC-Connected Block). When a DCCB is to be simulated, the entire observation period is first split into intervals. The length of the interval is specified by the user. Fast timing simulation is then carried for the gate in this interval. At the end of the interval, ILLIADS performs degradation calculation followed by an extrapolation. The average bond-breaking current is first computed, for each critical transistor in the gate, using the following equation:

$$<I_{BB}> = \frac{1}{T}\int_{t_s}^{t_e} I_{BB}dt, \tag{73}$$

where, t_s is when the interval starts and t_e ends. $T = t_e - t_s$. I_{BB} is defined in Eq. (69). Since V_{DS} can be piecewise linearized, $<I_{BB}>$ can be expressed by a closed form. To be specific, $<I_{BB}>$ can be expressed in a form:

$$<I_{BB}> = c_2(V_{DS} - V_{DSAT})^{\alpha+3} + c_1(V_{DS} - V_{DSAT})^{\alpha+2} + c_0(V_{DS} - V_{DSAT})^{\alpha+1}, \tag{74}$$

where, since V_{DS} is piecewise linearized, $V_{DS} - V_{DSAT}$ can be written as $\alpha_v(t - t_s) + \beta_v$. c_2, c_1, and c_0 are expressed in terms of the transistor parameters, α, α_v, and β_v. The extrapolation of the degradation level of each transistor is then done using Eq. (17).

Once the degradation level is found, the parameters of each transistor are then updated. The process is illustrated in Fig. 25.

There are several advantages of doing reliability simulation using ILLIADS. First of all, the speedup is tremendous. For a circuit as small as a 3-stage CMOS inverter chain, the run time of ILLIADS on a SUN-4 Sparcstation was only 0.1 second while the circuit level reliability simulation using iSMILE [2] took 3135.5 seconds. The speedup is already 4 orders of magnitude even for a small circuit and increases with the circuit size.

Secondly, since I_{BB} is a pulse waveform, the choice of T is crucial. A large variation of T can still result in the same $<I_{BB}>$ for one cycle simulation. This results in large error of the degradation level. The remedy to this problem is to choose T large enough, for example, 10 cycles. For the same 3-stage CMOS inverter chain, the run time can immediately increase to 30,000 seconds for a circuit level reliability simulation, which is rather impractical. Thirdly, due to the same reason, for large circuits, for example, a 32-bit CMOS ALU, dynamic reliability simulation at the circuit level is formidable, if not impossible. In comparison, ILLIADS takes only about 4 minutes to simulate the circuit for 51 cycles. Lastly, using the dynamic reliability simulation approach, the most damaged gates inside a circuit can be pinpointed.

Figures 26 and 27 shows the reliability simulation results for a 3-stage CMOS inverter chain. As can be seen, the output waveform of ILLIADS matches that of iSMILE. Compared with the complicated model used in iSMILE [33], this also shows that the simple MOS transistor model is accurate. The departure of the degradation curves is within 10%. Notice that the departure of the degradation curves may be due to the simple transistor model since its maximum error is around 10%.

Figure 28 shows the cumulative degradation distribution of a 32-bit CMOS ALU. The circuit consists of 3786 transistors (1038 logic gates). Run time for this circuit was 227 seconds. Notice that, if one cycle simulation was done, the run time will be only 4.4 seconds. The same circuit was tested with 10 different random input patterns. The deviations in distribution profile are almost negligible. For the ALU circuit tested, ILLIADS pinpointed that the most damaged gates are the ones which produce the ALU outputs. A damage histogram can also be generated using ILLIADS. Figure 29 illustrates the statistical distribution of degradation levels in the 32-bit ALU circuit for operational stress periods ranging from 3 months to 15 months. It can be seen that the percentage peak of shifts to higher

degradation levels with increasing stress time, while the damage level tends to spread more broadly among the devices in the circuit.

5.3 Application

ILLIADS has been successfully installed and demonstrated in AMD(Advanced Micro Devices) in Sunnyvale, CA. Future technology transfer of this program is in progress.

6 Summary

During the 1992-1993 project period, we have developed parametric macromodels for hot-carrier resistant CMOS circuits. The macromodel approach is for the parametric evaluation of hot-carrier-induced degradation of nMOS transistors operating in digital logic circuit environment. The formulation of the macromodel helps to identify a small set of relevant device and circuit parameters which influence the circuit reliability. It is shown that for a wide class of CMOS and nMOS circuits, the dynamic degradation due to hot-carrier effects can be expressed as a function of i) the nMOS transistor channel width W_n, ii) the load capacitance C_L, and iii) the input-voltage rising slope a. Using the macromodel analysis, simple design-for-reliability rules based on layout geometry are proposed for various nMOS/CMOS circuits, and for various operating conditions. The analysis results obtained for widely used circuit configurations agree with recently published experimental findings.

The parametric design-for-reliability rule devised in this work is then applied for geometry-based, early diagnosis of potential reliability problems in CMOS circuits. iRULE program is developed for this purpose. We have demonstrated that iRULE is capable of analyzing very large circuits with computational efficiency. It can be used effectively as a front-end tool in a hierarchical design-for-reliability system. iRULE reads in circuit netlist and uses a

table look-up method to identify critical hot-carrier damaged transistors. iRULE has been successfully technology transferred to AT&T Bell Labs in Allentown, PA., for hot-carrier A.C. lifetime estimation for their 0.9μm and 0.5μm technology.

We have also designed reliability tester chips for the verification of the developed macro-model. The test structures are so designed that the hot-carrier damage could be measured as functions of the circuit parameters. The chips has been manufactured by MOSIS using CMOS n-well 1.2μm technology. Proposed measurements will be carried out in the future.

The fast timing simulation tool ILLIADS has been applied to reliability simulation. We have proposed a simple six parameter model for nMOS hot-carrier degradation. The simulation is fast and accurate. In the future, we plan to combine our reliability diagnosis tool iRULE, fast timing simulation tool ILLIADS, and circuit level simulation tool iSMILE into a design-for-reliability system software. It should be able to handle the reliability analysis of very large circuits both for early diagnosis and for detailed transistor level simulation.

References

[1] M. M. Kuo, K. Seki, P. M. Lee, J. Y. Choi, P. K. Ko, and C. Hu, "Simulation of MOSFET lifetime under ac hot-electron stress," *IEEE Transactions on Electron Devices*, vol. ED-35, pp. 1004–1011, July 1988.

[2] A. T. Yang and S. M. Kang, "A novel circuit simulation program with emphasis on new device model development," in *Proceeding of the ACM/IEEE Design Automation Conference*, pp. 630–633, June 1989.

[3] Y. Leblebici and S. M. Kang, "An integrated hot-carrier degradation simulator for VLSI reliability analysis," in *Proceeding of the IEEE International Conference on Computer-Aided Design*, pp. 400–403, November 1990.

[4] W. Hsu, B. Sheu, and S. Gowda, "Design of reliable VLSI circuits using simulation techniques," *IEEE Journal of Solid-State Circuits*, vol. 26, pp. 452–457, March 1991.

[5] Y. Shih, Y. Leblebici, and S. Kang, "New simulation methods for MOS VLSI timing and reliability," *Proceeding of the IEEE International Conference on Computer-Aided Design*, pp. 162–165, November 1991.

[6] R. Tu, E. Rosenbaum, W. Chan, C. Li, E.Minami, K. Quader, P. Ko, and C. Hu, "Berkeley Reliability Tools–BERT," *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*, vol. 12, pp. 1524–1534, October 1993.

[7] W. Sun, Y. Leblebici, and S. M. Kang, "A parametric macro-model approach for hot-carrier resistant MOS VLSI design," in *Proceeding of the IEEE Custom Integrated Circuits Conference*, pp. 18.3.1–18.3.4, May 1992.

[8] Y. Leblebici, W. Sun, and S. Kang, "Parametric macro-modeling of hot-carrier induced dynamic degradation in MOS VLSI circuits," *IEEE Transactions on Electron Devices*, vol. 40, pp. 673–676, March 1993.

[9] C. Hu, S. Tam, F. Hsu, P. Ko, T. Chan, and K. Terrill, "Hot-electron-induced MOSFET degradation - model, monitor and improvement," *IEEE Transactions on Electron Devices*, vol. ED-32, pp. 375–384, February 1985.

[10] M.Brox and W. Weber, "Proof and quantification of dynamic effects in hot-carrier-induced degradation under pulsed operation conditions," *Proceeding of IEEE International Reliability Physics Symposium*, pp. 133–141, February 1991.

[11] W. Weber, M. Bronx, T. Kuenemund, H. Muehlhoff, and D. Schmitt-Landsiedel, "Dynamic degradation in MOSFETs—Part II: application in the circuit environment," *IEEE Transactions on Electron Devices*, vol. ED-38, pp. 1859–1867, August 1991.

[12] A. Schwerin, W. Hansch, and W. Weber, "The relationship between oxide charge and device degradation: A comparative study of n- and p- channel MOSFET's," *IEEE Transactions on Electron Devices*, vol. ED-34, pp. 2493–2499, December 1987.

[13] P. Heremans, R. Bellens, G. Groeseneken, and H. Maes, "Consistent model for the hot-carrier degradation in n-channel and p-channel MOSFET's," *IEEE Transactions on Electron Devices*, vol. ED-35, pp. 2194–2209, December 1988.

[14] B. Doyle, M. Bourderie, J. Marchetaux, and A. Boudou, "Interface state creation and charge trapping in the medium-to-high gate voltage range ($Vd/2 > Vg > Vd$) during hot-carrier stress of n-MOS transistors," *IEEE Transactions on Electron Devices*, vol. ED-37, pp 744–754, March 1990.

[15] P. Heremans, G. van den Bosch, R. Bellens, G. Groeseneken, and H. Maes, "Temperature dependence of the channel hot-carrier degradation of n-channel MOSFET's," *IEEE Transactions on Electron Devices*, vol. ED-37, pp. 980–993, April 1990.

[16] S. Aur, D. Hocevar, and P. Yang, "HOTRON - A circuit hot electron effect simulator," in *Proceeding of the IEEE International Conference on Computer-Aided Design*, pp. 256–259, November 1987.

[17] E. Takeda and N. Suzuki, "An empirical model for device degradation due to hot-carrier injection," *IEEE Electron Device Letters*, vol. EDL-4, pp. 111–113, April 1983.

[18] P. Ko, C. Hu, and R. Muller, "A unified model for hot electron currents in MOSFET's," *IEEE International Electron Devices Meeting Technical Digest*, pp. 600–603, December 1981.

[19] Y. ElMansy and A. Boothroyed, "A simple two-dimensional model for IGFET operation in the saturation region," *IEEE Transactions on Electron Devices*, vol. ED-24, pp. 254–262, March 1977.

[20] S. Tanaka and M. Ishikawa, "One-dimensional writing model of n-channel floating gate-ionization MOS(FIMOS)," *IEEE Transactions on Electron Devices*, vol. ED-28, pp. 1190–1197, October 1981.

[21] T. Y. Chan, P. K. Ko, and C. Hu, "A simple method to characterize substrate current in MOSFET's," *IEEE Electron Device Letters*, vol. EDL-5, pp. 505–507, December 1984.

[22] E. H. Nicollian and J. R. Brews, *MOS Physics and Technology.* New York: Wiley, 1983.

[23] W. Weber, C. Werner, and A. Schwerin, "Lifetimes and substrate currents in static and dynamic hot-carrier degradation," *IEEE International Electron Devices Meeting Technical Digest*, pp. 390–393, December 1986.

[24] W. Weber, "Dynamic stress experiments for understanding hot-carrier degradation phenomena," *IEEE Transactions on Electron Devices*, vol. ED-35, pp. 1476–1486, September 1988.

[25] C. T. Sah, "VLSI device reliability modeling," in *Proceedings of 1987 International Symposium on VLSI Technology, Systems and Applications*, pp. 153–162, May 1987.

[26] Y. Leblebici, *Modeling and simulation of hot-carrier induced device and circuit degradation for VLSI reliability.* PhD thesis, University of Illinois, Urbana-Champaign, 1990.

[27] S. M. Kang and H. Y. Chen, "A global delay model for domino CMOS circuits with application to transistor sizing," *International Journal of Circuit Theory and Applications*, vol. 18, pp. 289–306, May 1990.

[28] J. P. Uyemura, *Fundamentals of MOS Digital Integrated Circuits.* New York: Addison-Wesley Publishing Company, 1983.

[29] D. A. Hodges and H. G. Jackson, *Analysis and Design of Digital Integrated Circuits.* New York: McGraw-Hill Book Company, 1988.

[30] J. S. Suehle and K. Galloway, "Test circuit structures for characterizing the effects of localized hot-carrier-induced charge in VLSI switching circuits," *IEEE Proceedings on Microelectronic Test Structures*, vol. 1, pp. 126–131, February 1988.

[31] D. Cho and S. Kang, "An accurate charge-oriented table-lookup model for circuit simulation," *Proceeding of the IEEE International Symposium on Circuits and Systems*, pp. 1531–1534, May 1993.

[32] Y. Leblebici and S. M. Kang, "An integrated reliability simulation tool for hot-carrier resistant VLSI design," in *Proceeding of TECHCON Conference*, pp. 102–105, October 1990.

[33] Y. Leblebici and S. M. Kang, "A one-dimensional MOSFET model for simulation of hot-carrier induced device and circuit degradation," in *Proceeding of the IEEE International Symposium on Circuits and Systems*, pp. 109–112, May 1990.

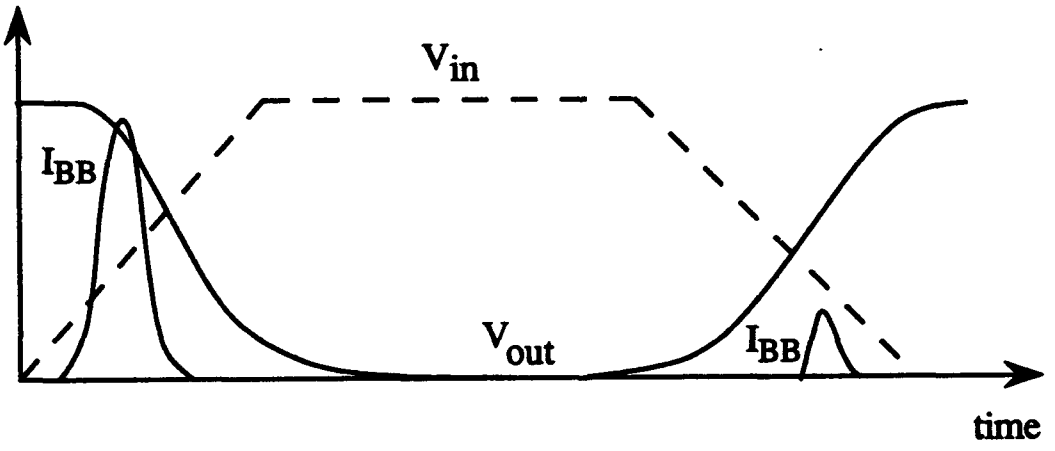


Figure 1: Hot-carrier damage for nMOS transistor during rising and falling input.

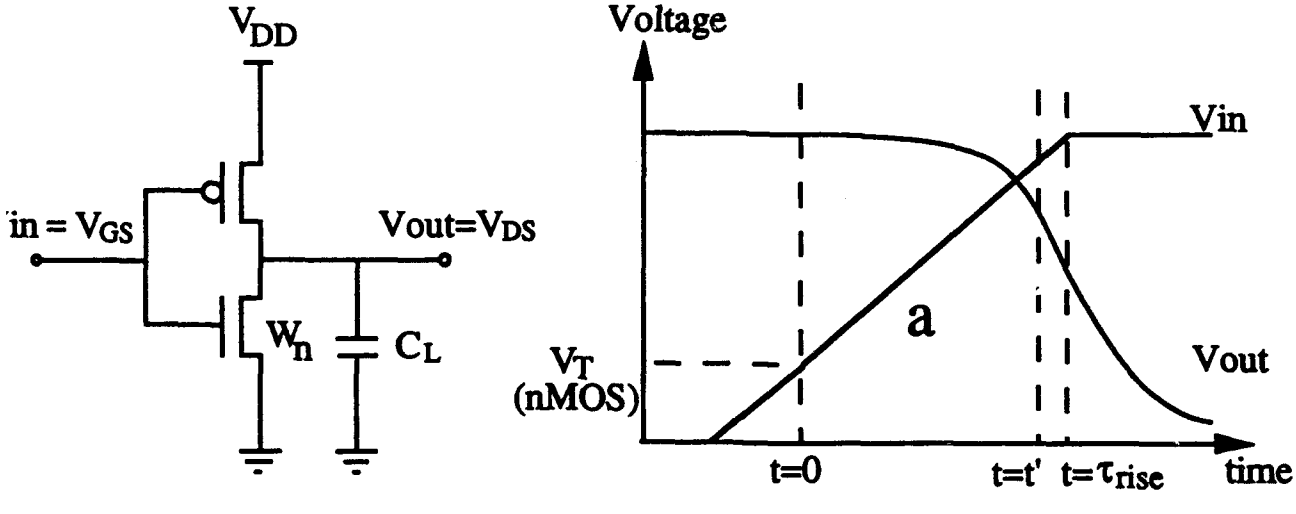


Figure 2: CMOS inverter circuit and typical input/output voltage waveforms.

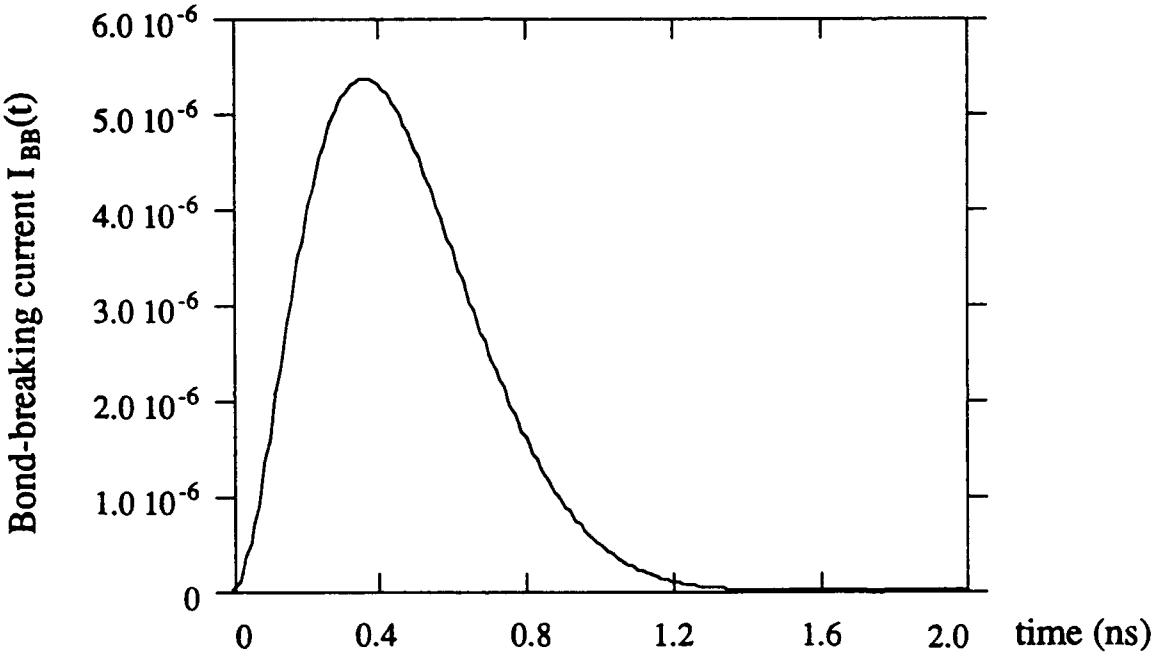


Figure 3: Bond-breaking current as a function of time t.

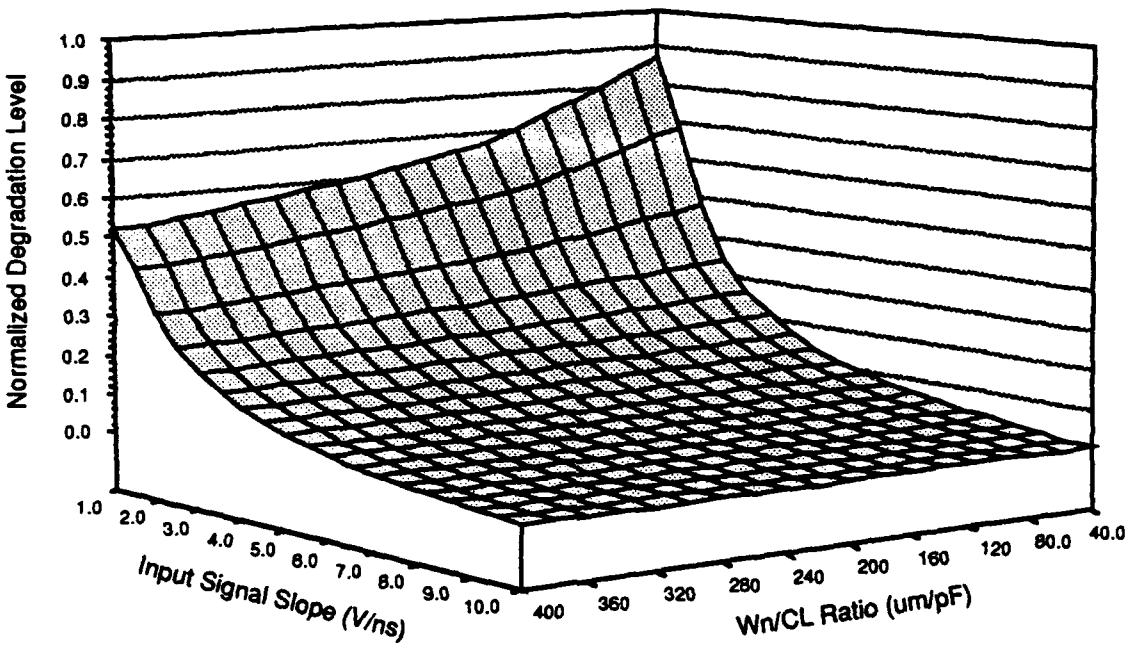


Figure 4: CMOS inverter degradation as a function of input signal slope a and W_n/C_L.

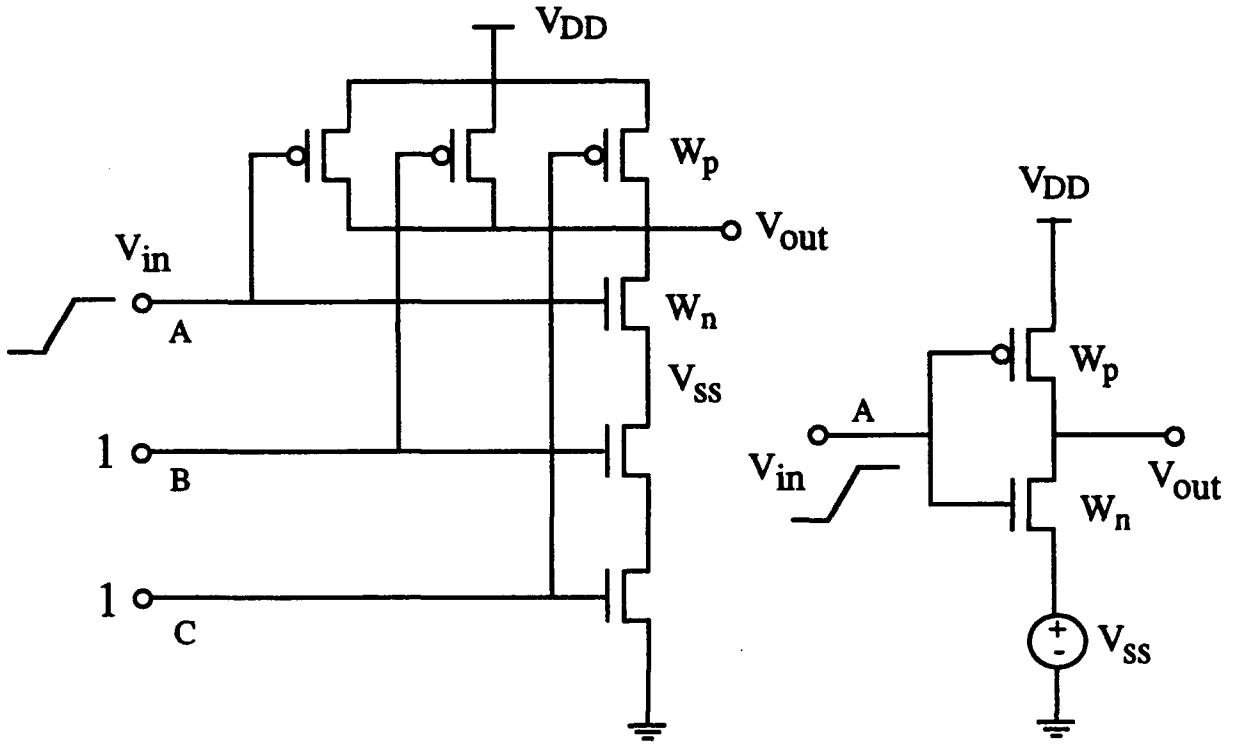


Figure 5: CMOS equivalent inverter circuit for CMOS NAND Gates.

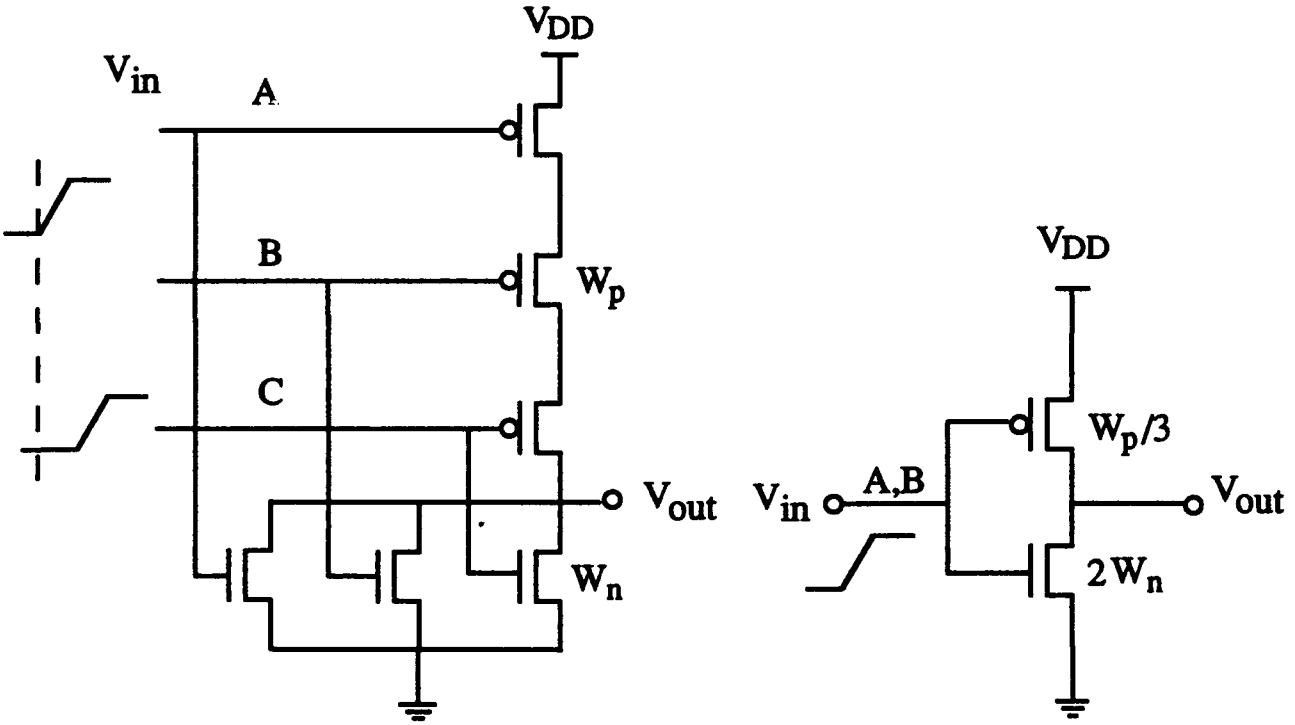


Figure 6: CMOS equivalent inverter circuit for CMOS NOR Gates.

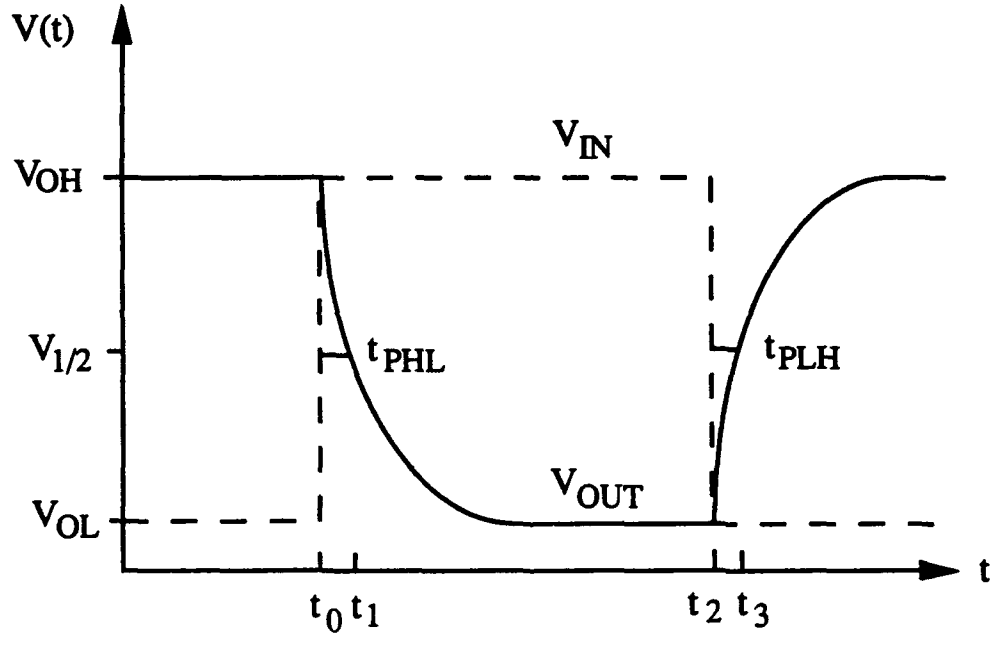


Figure 7: Propagation delay time definitions.

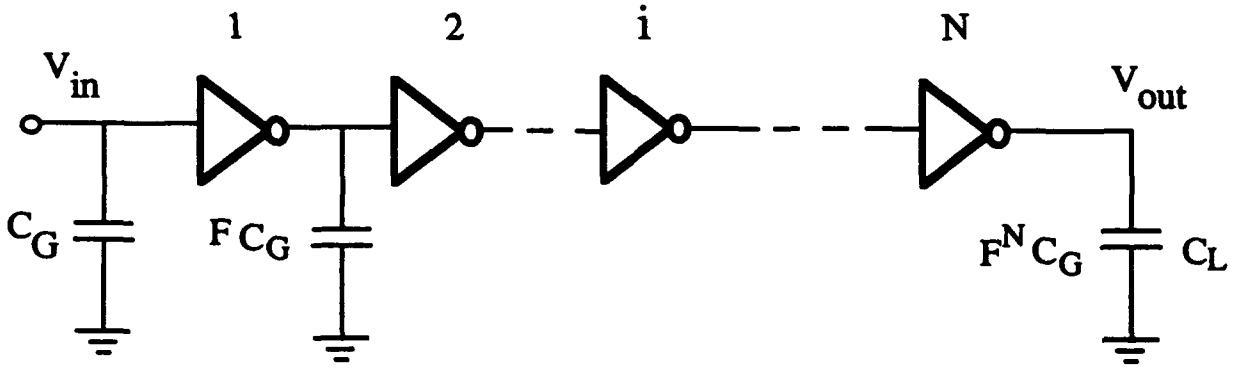


Figure 8: Scaled CMOS inverter chain.

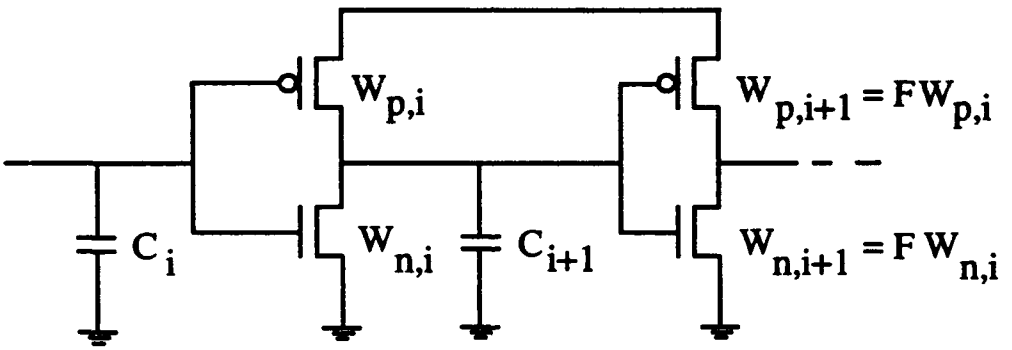


Figure 9: Scaled inverter chain at stage i.

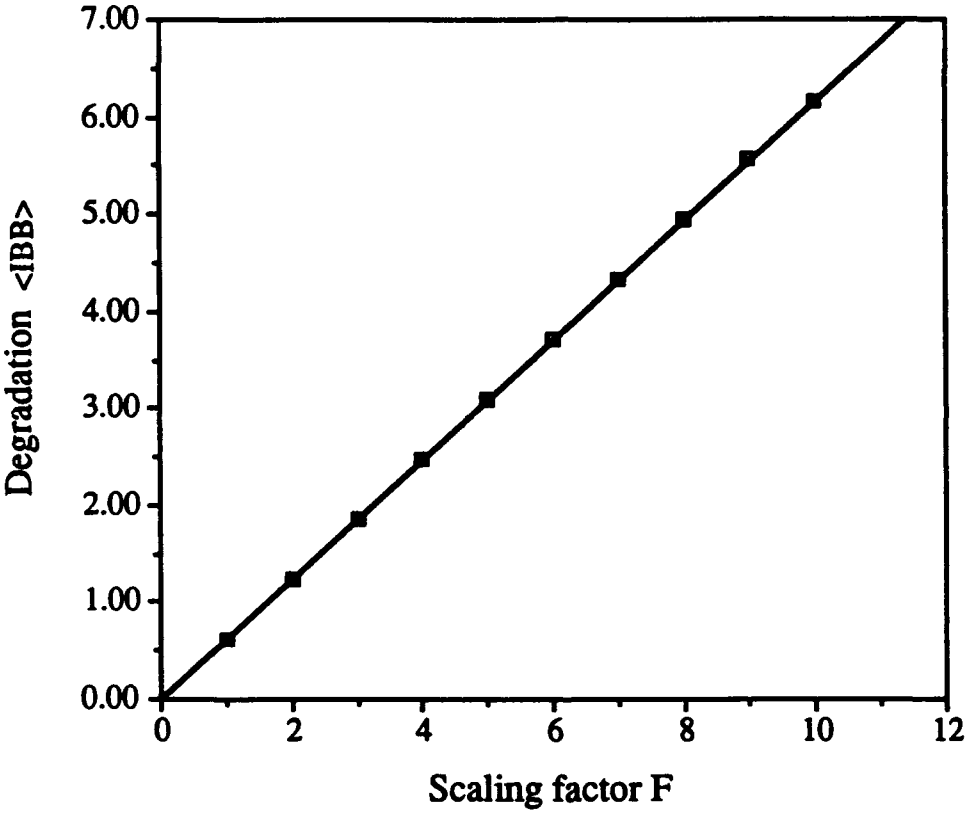


Figure 10: Degradation as a function of scaling factor F.

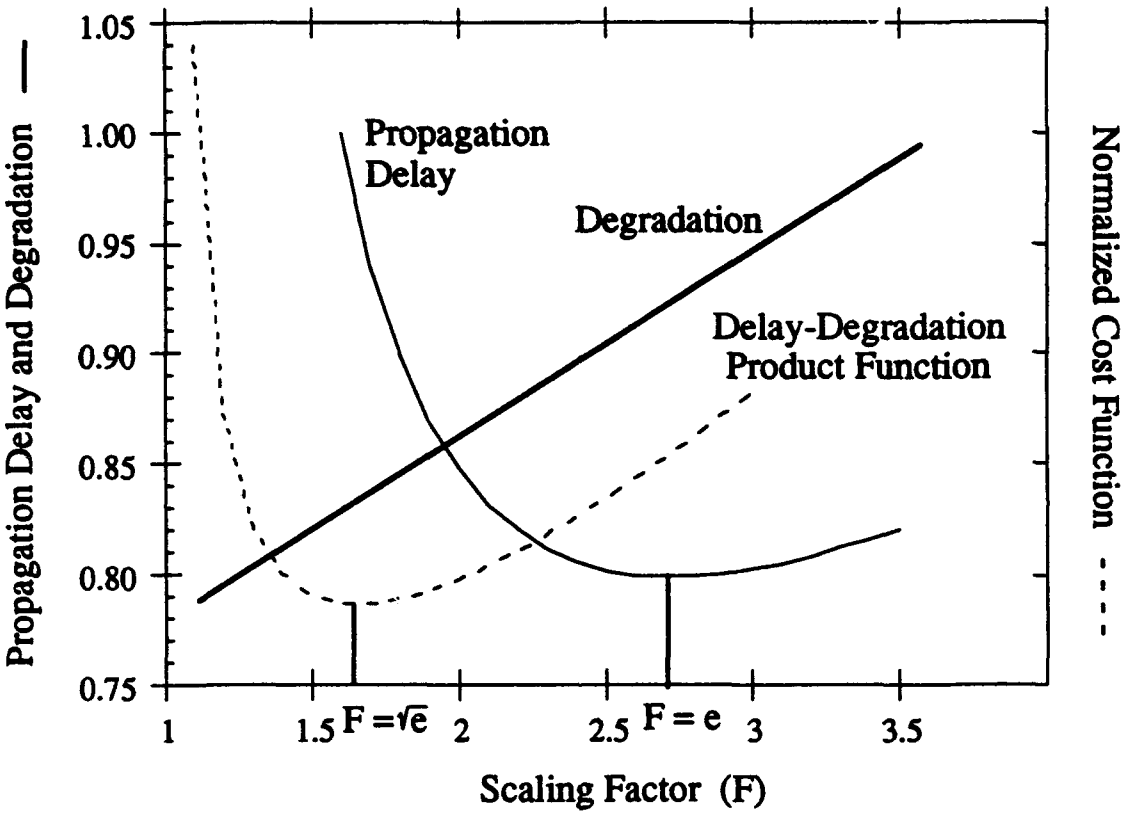


Figure 11: Normalized degradation and propagation delay versus scaling factor F and optimization of cost function versus scaling factor F.

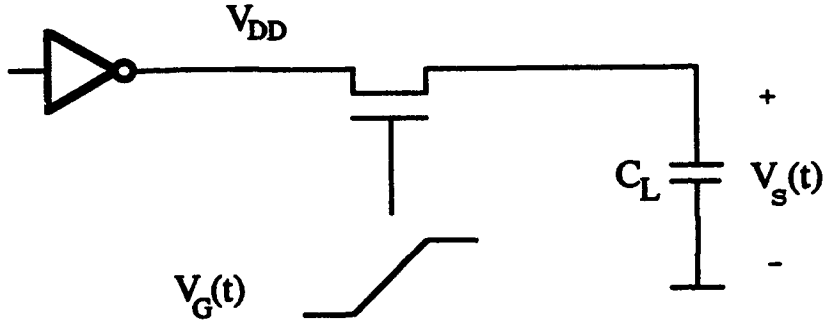


Figure 12: The nMOS transmission gate circuit.

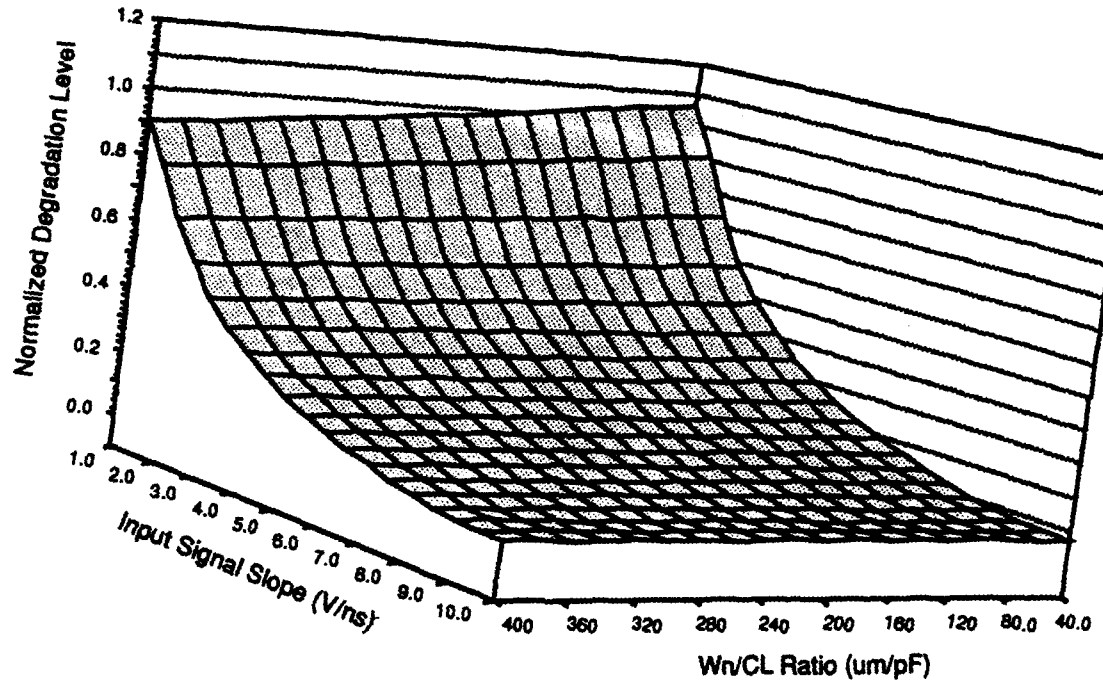


Figure 13: Degradation of nMOS pass-gate transistor as a function of input signal slope a and W_n/C_L.

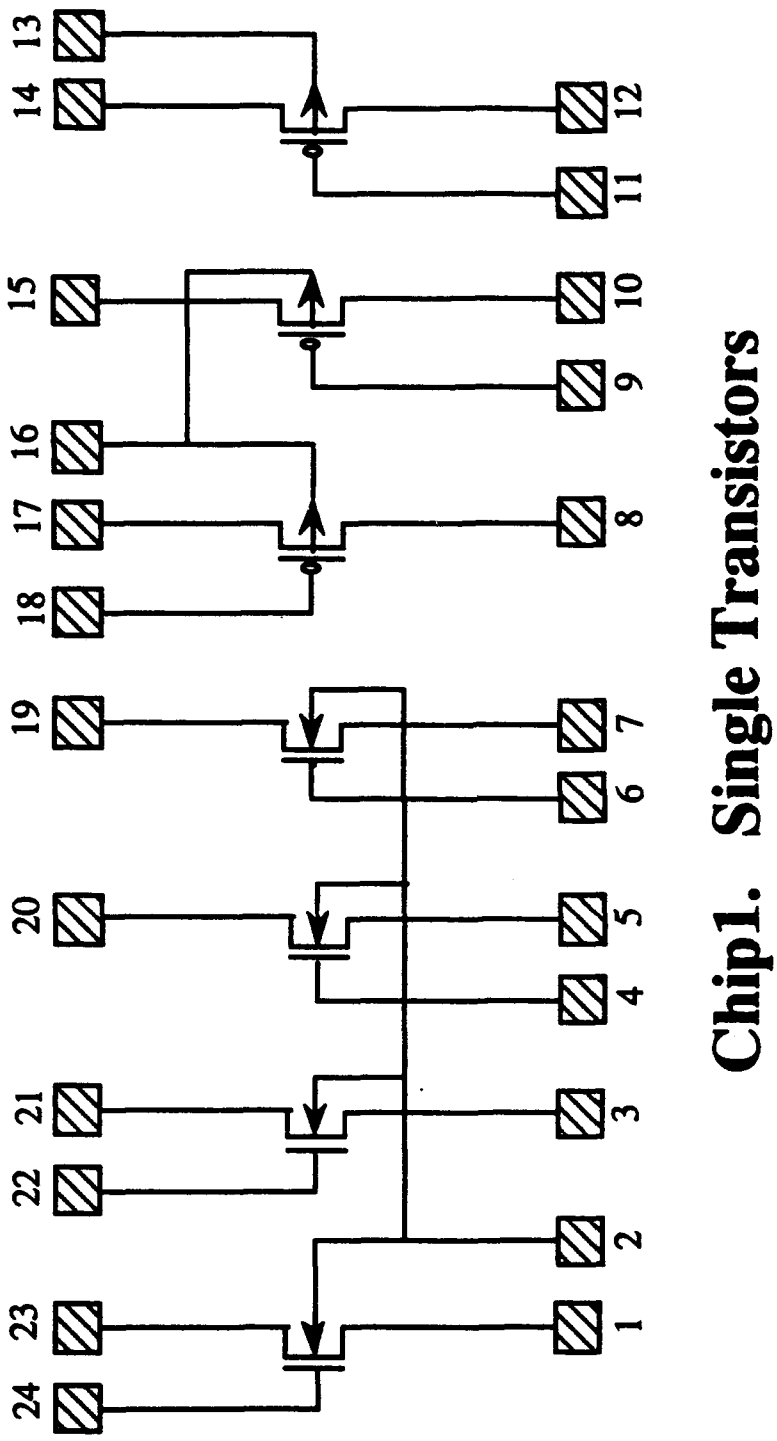
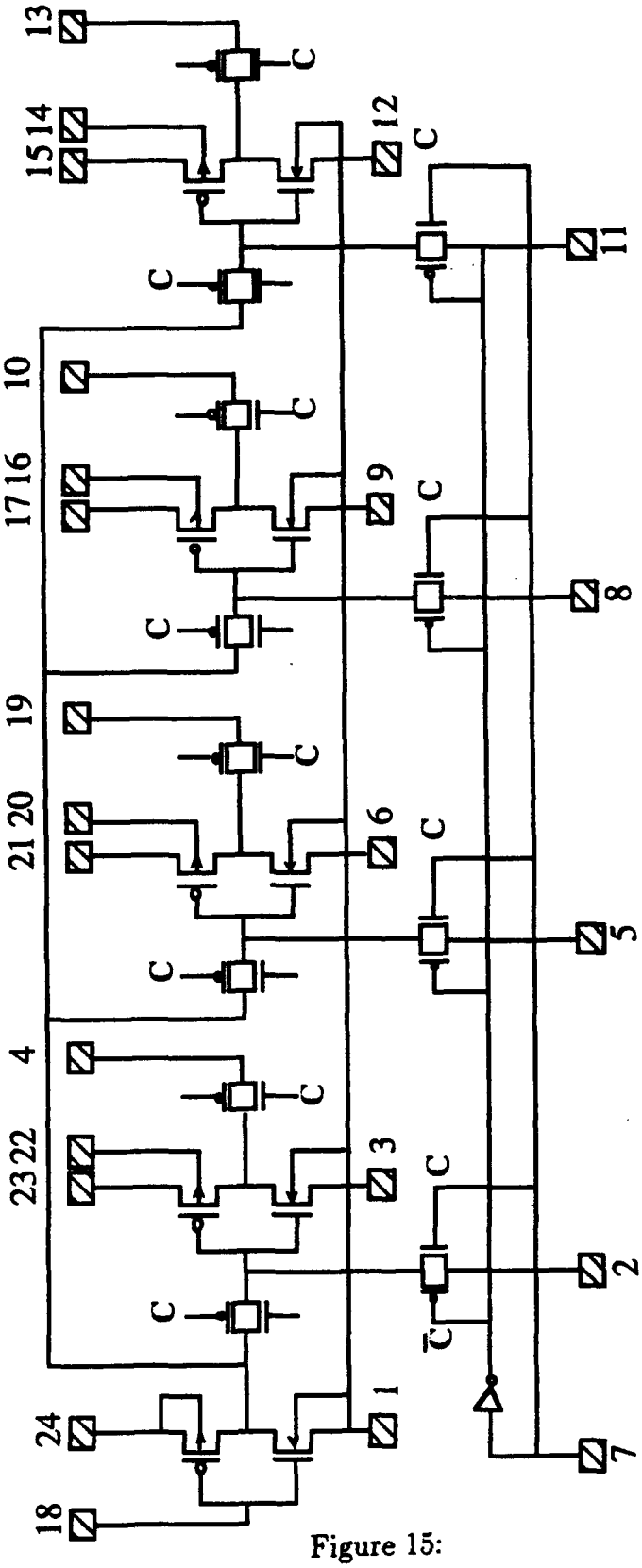
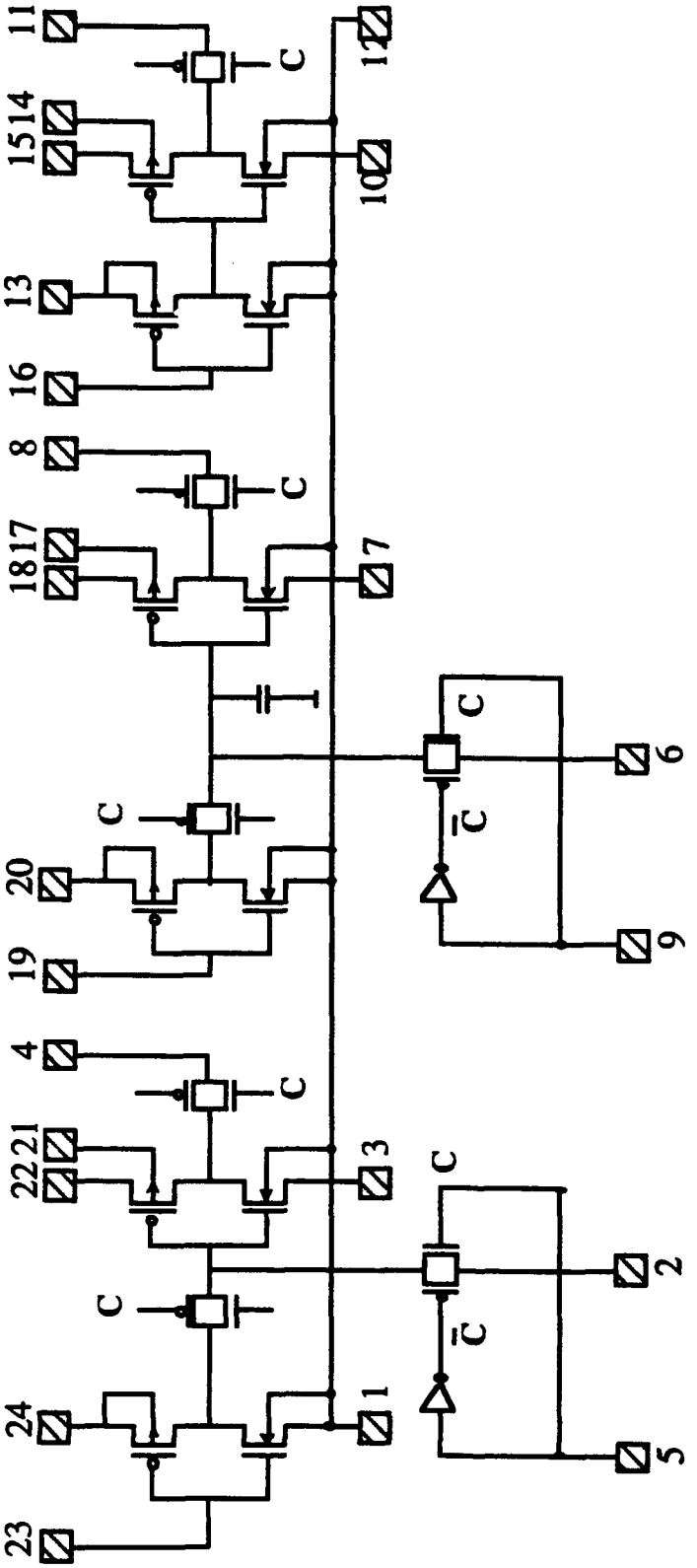

Chip1. Single Transistors
13
14
12
11
15
10
9
16
17
8
18
19
7
6
20
5
4
21
3
22
2
23
1
24

Figure 14:



Chip2. CMOS inverters with same input and different sizes

Figure 15:



Chip 3. CMOS inverters of the same size and different input slopes

Figure 16:

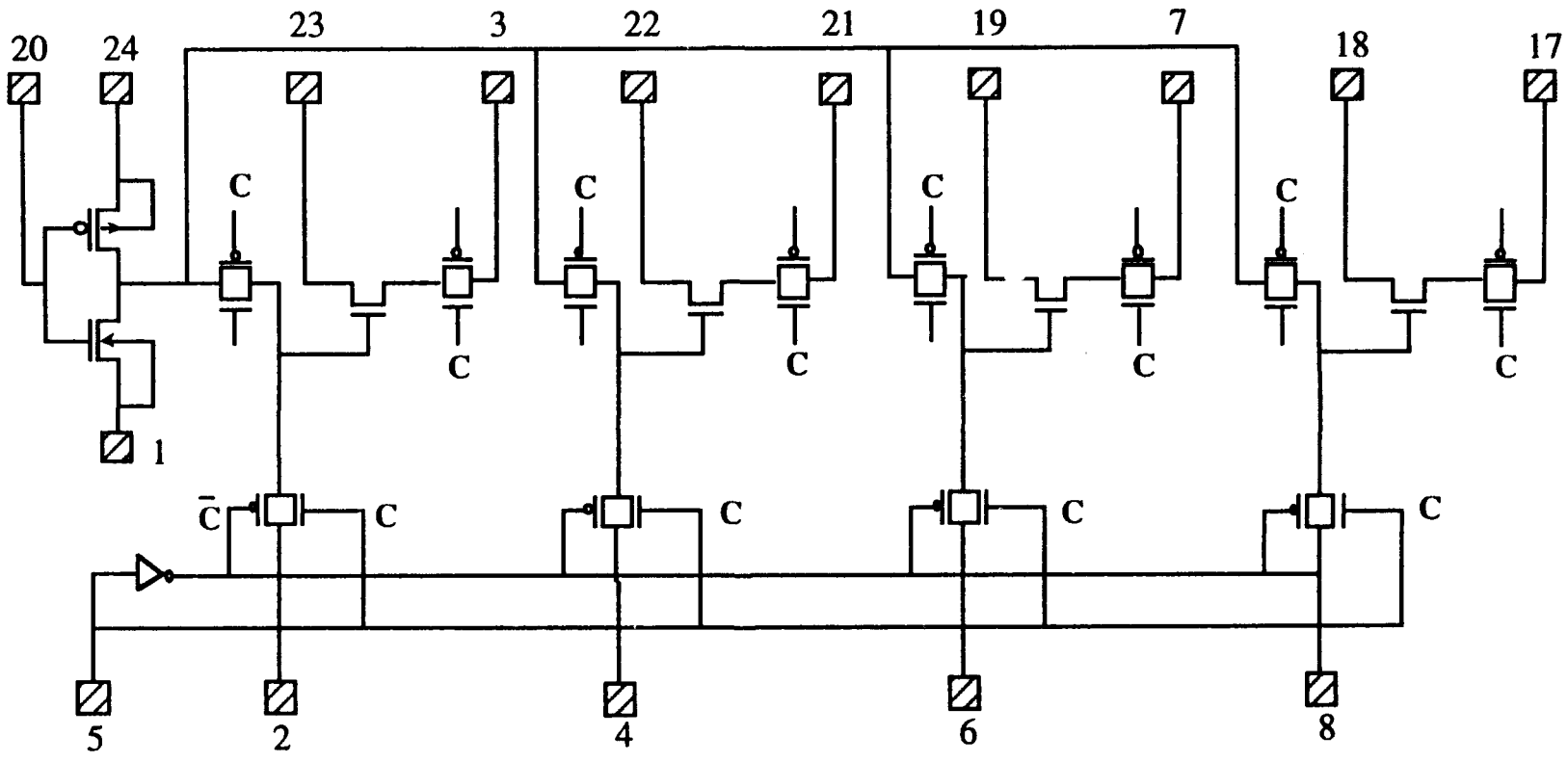


Figure 17:

Chip4. CMOS inverters with same input and different sizes

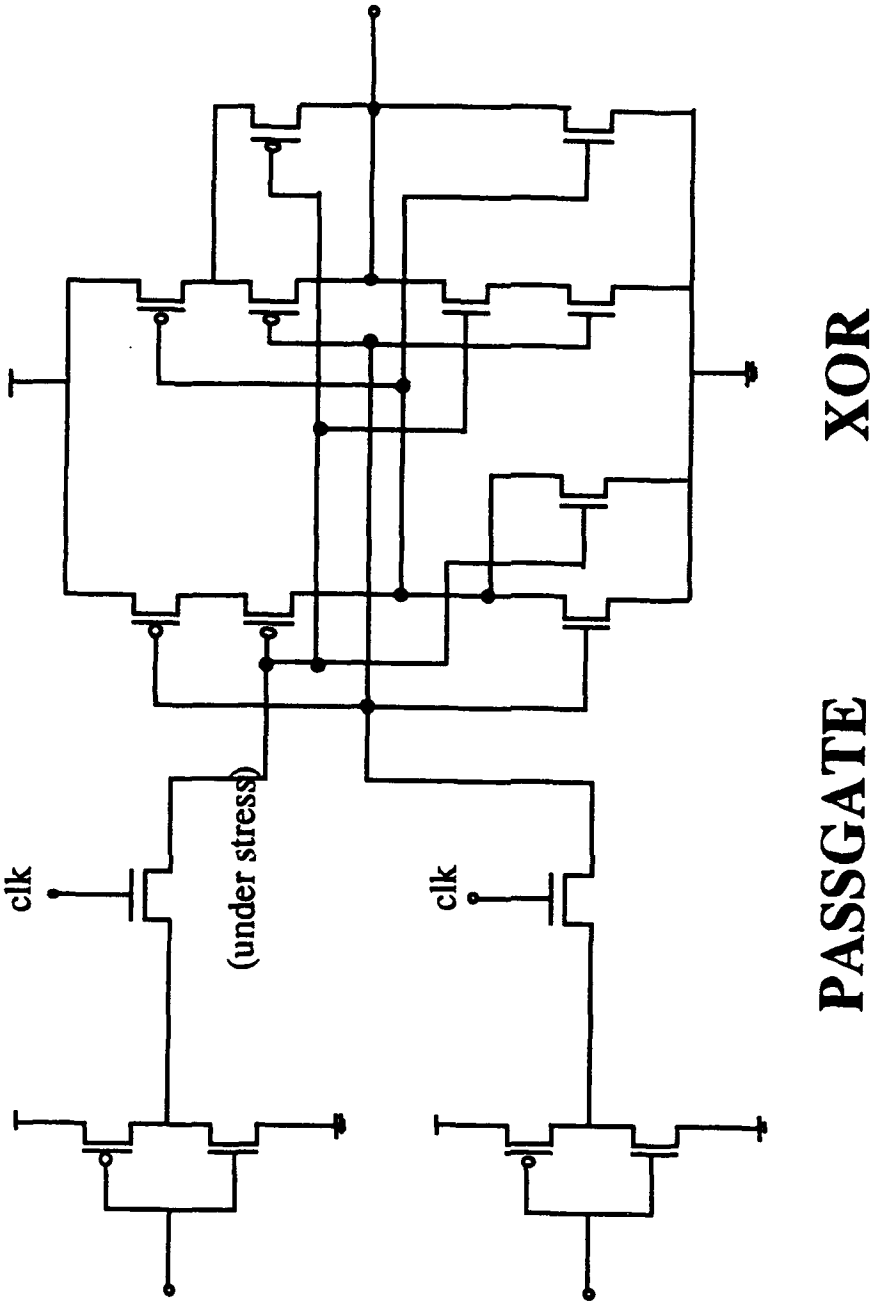


Figure 18:

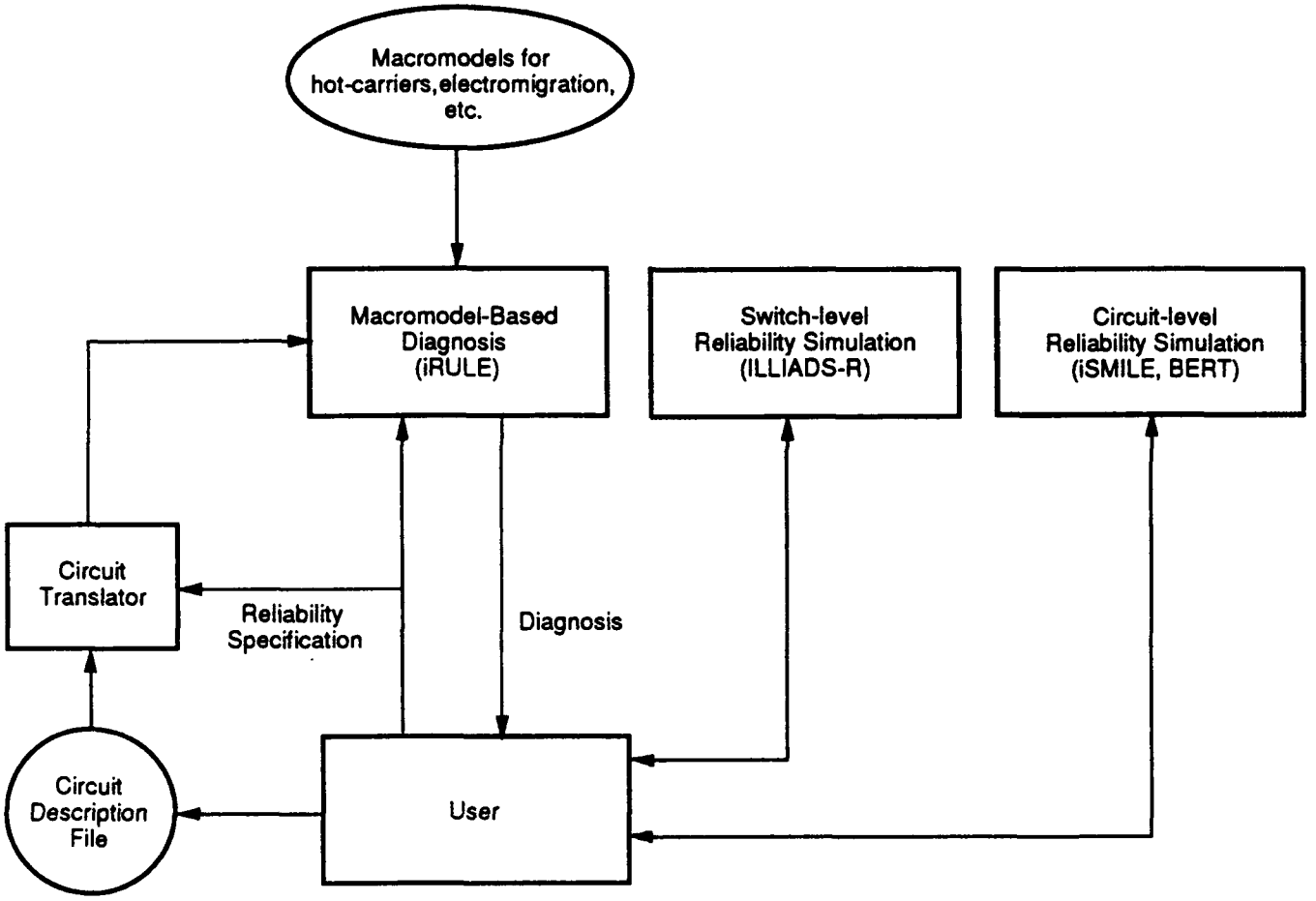


Figure 19: Framework for the macromodel-based reliability tool iRULE and for hierarchical reliability estimation of VLSI circuits.

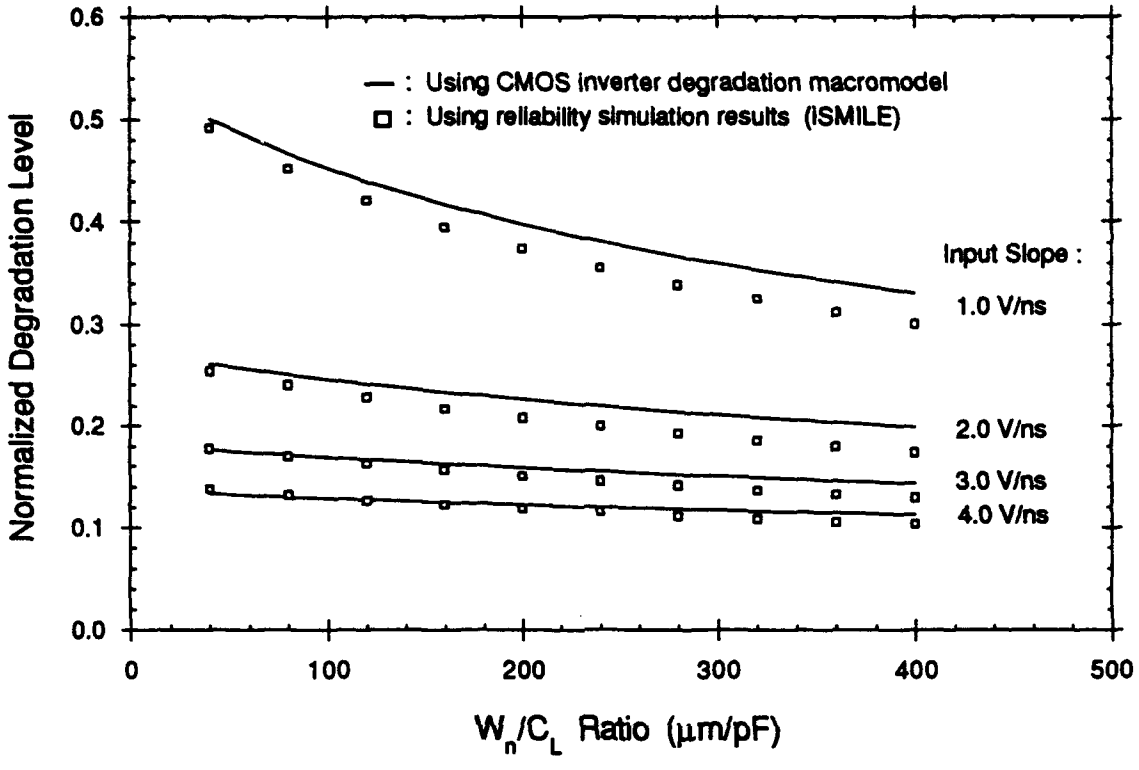


Figure 20: Comparison of macro-model results with circuit-level reliability simulation results, obtained by using iSMILE.

```
begin
    read-in circuit description file;
    create data-structures & internal representation of circuit;
    do (for all logic gates) {
        distinguish between Source & Drain of each pMOS transistor;
        determine worst-case (W/L)_equiv of each pMOS block;
        compute effective load capacitance C_L;
        compute rising signal slope at output;
    }
    read-in hot-carrier degradation look-up table(s);
    do (for all uppermost nMOS transistors) {
        extract the 2 device parameters serving as indices into look-up table;
        if (within the range of look-up table) {
            interpolate into table & compute ⟨I_BB⟩;
        } else {
            report: out of table's range;
        }
    }
    compute mean ⟨I_BB⟩ averaged over all critical transistors;
end
```

Figure 21: Algorithm

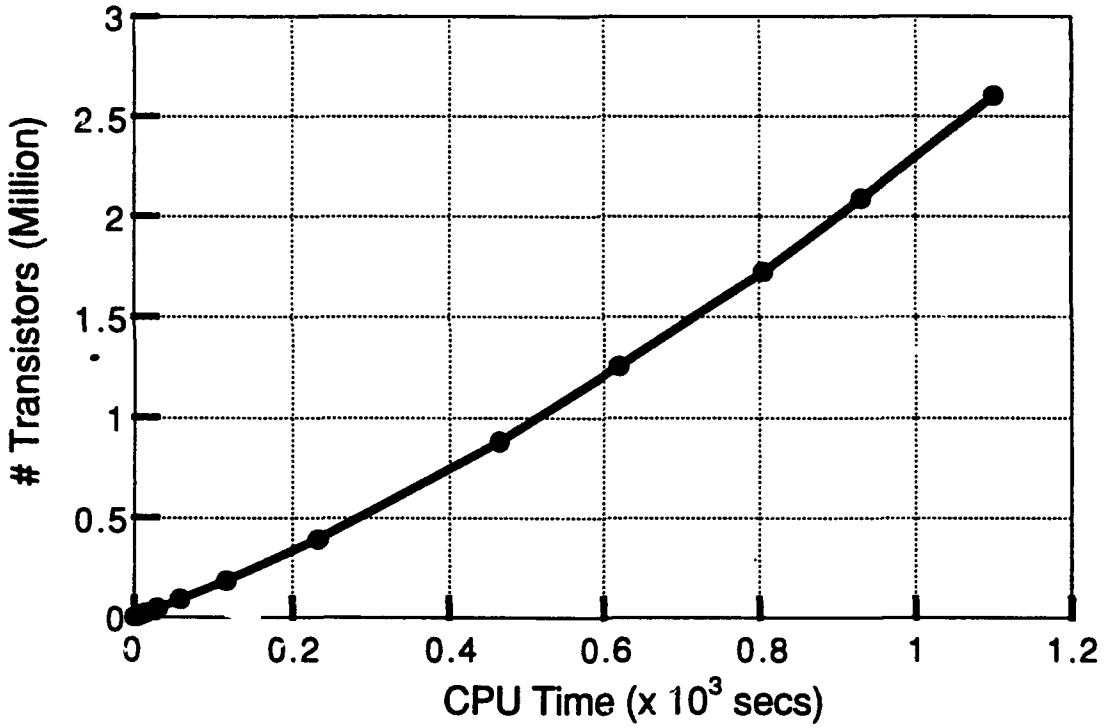


Figure 22: iRULE performance vs. number of transistors.

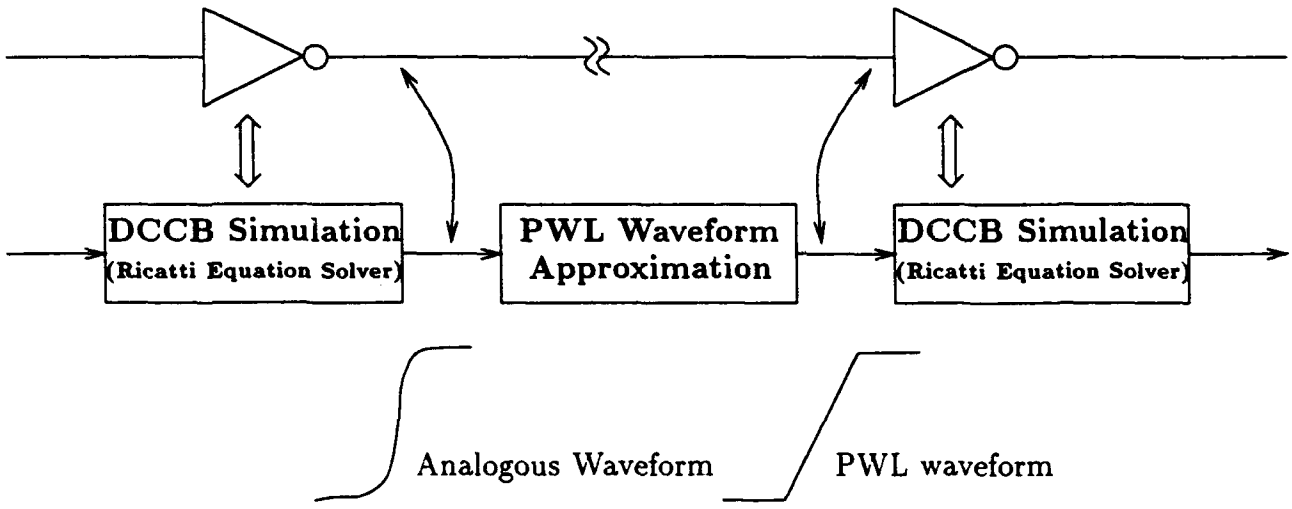


Figure 23: A Typical ILLIADS Simulation

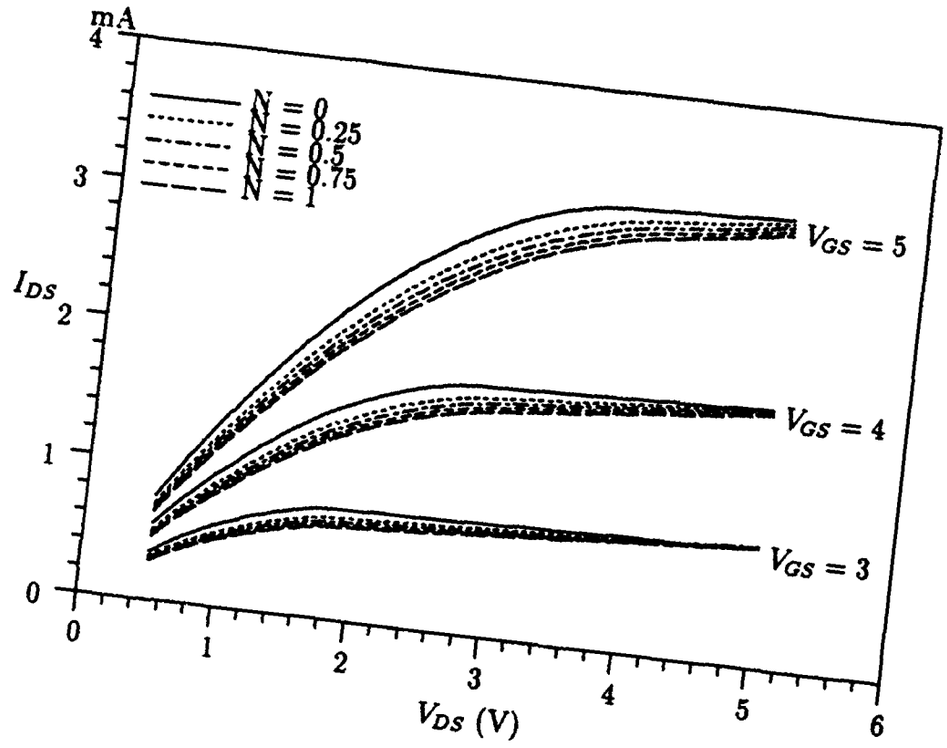


Figure 24: Comparison of the I–V curves before and after hot-carrier stress

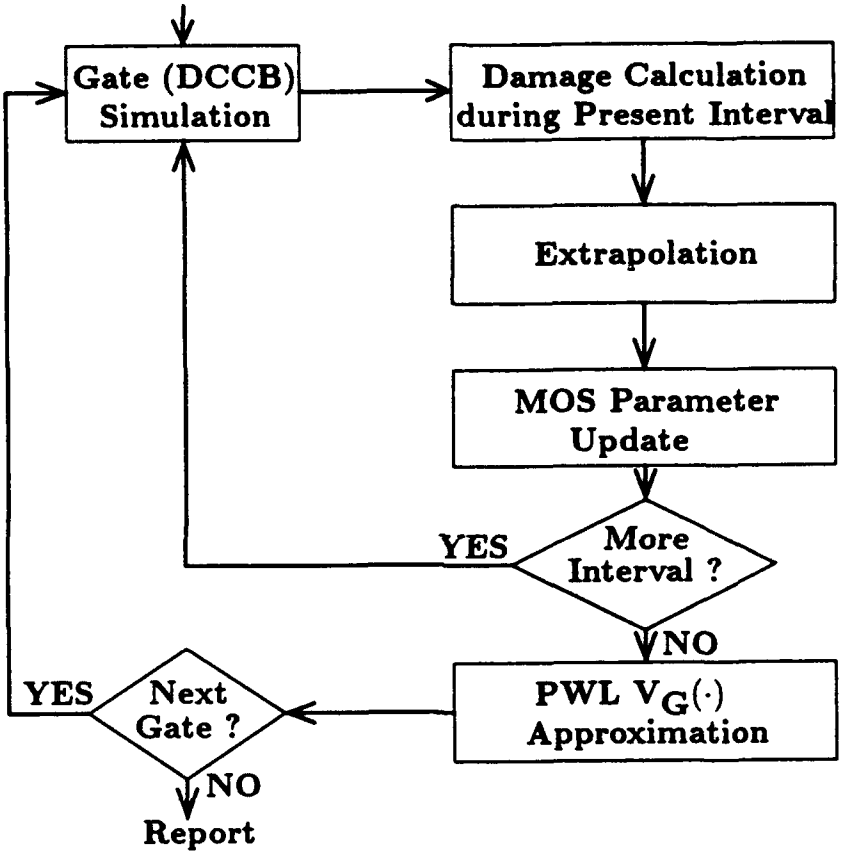


Figure 25: Process of ILLIADS Reliability Simulation

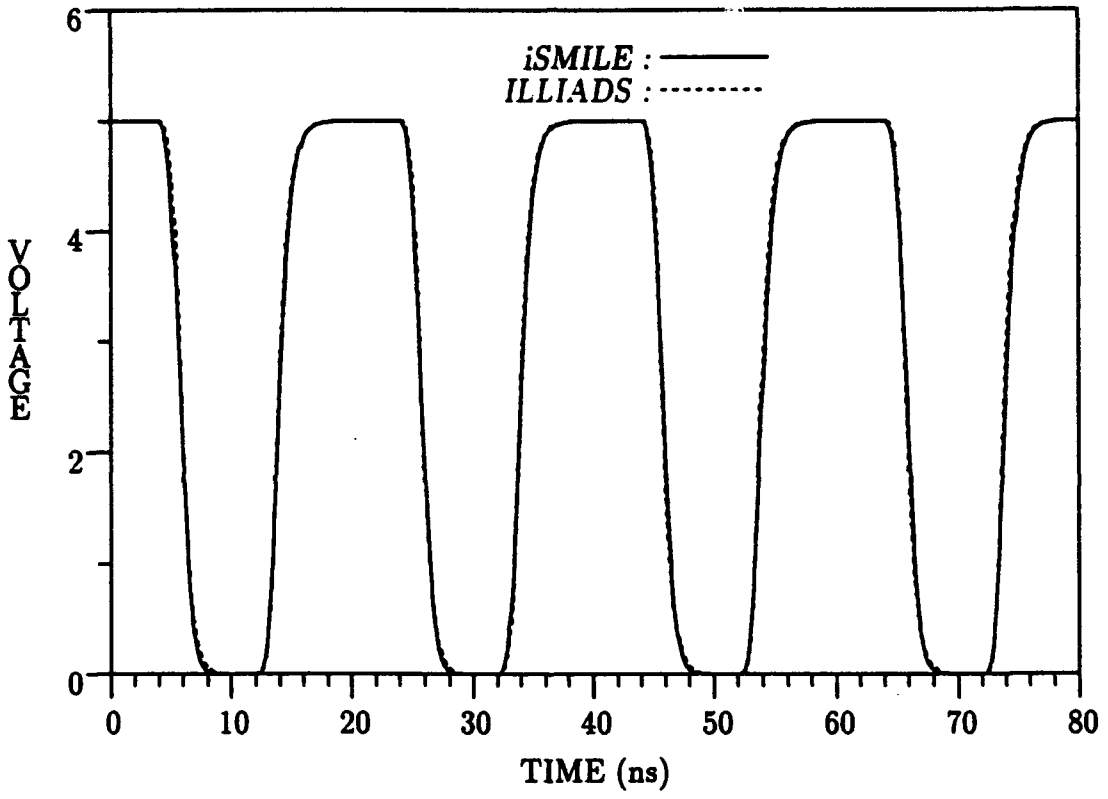


Figure 26: Output Waveforms at the Last Stage of a 3-stage Inverter Chain

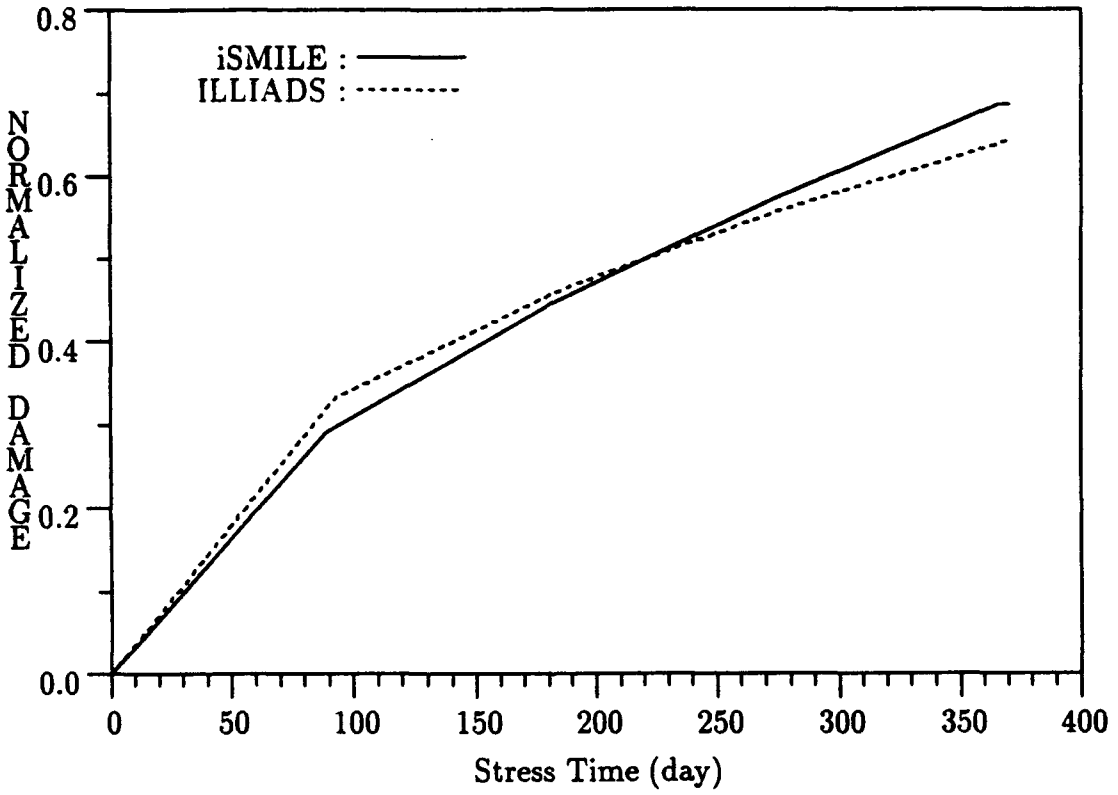


Figure 27: Damage of the nMOS of the Last Stage of a 3-stage Inverter Chain

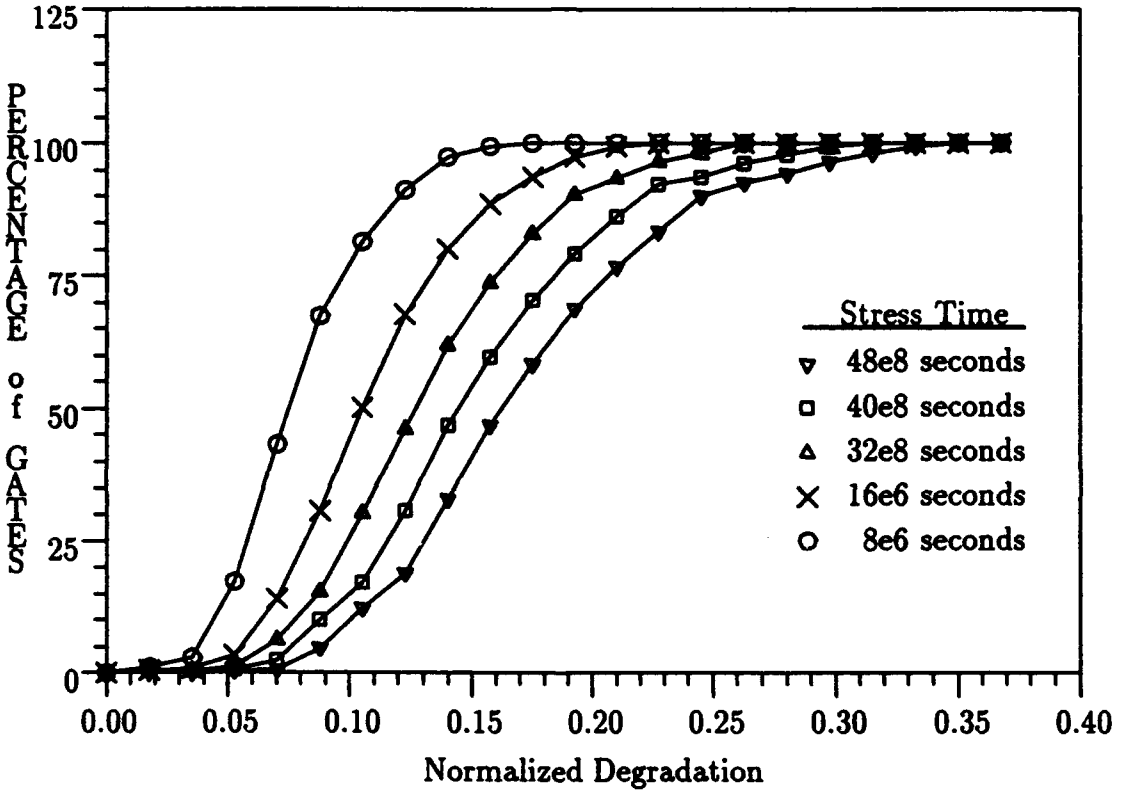


Figure 28: Cumulative Degradation Distribution of ALU32

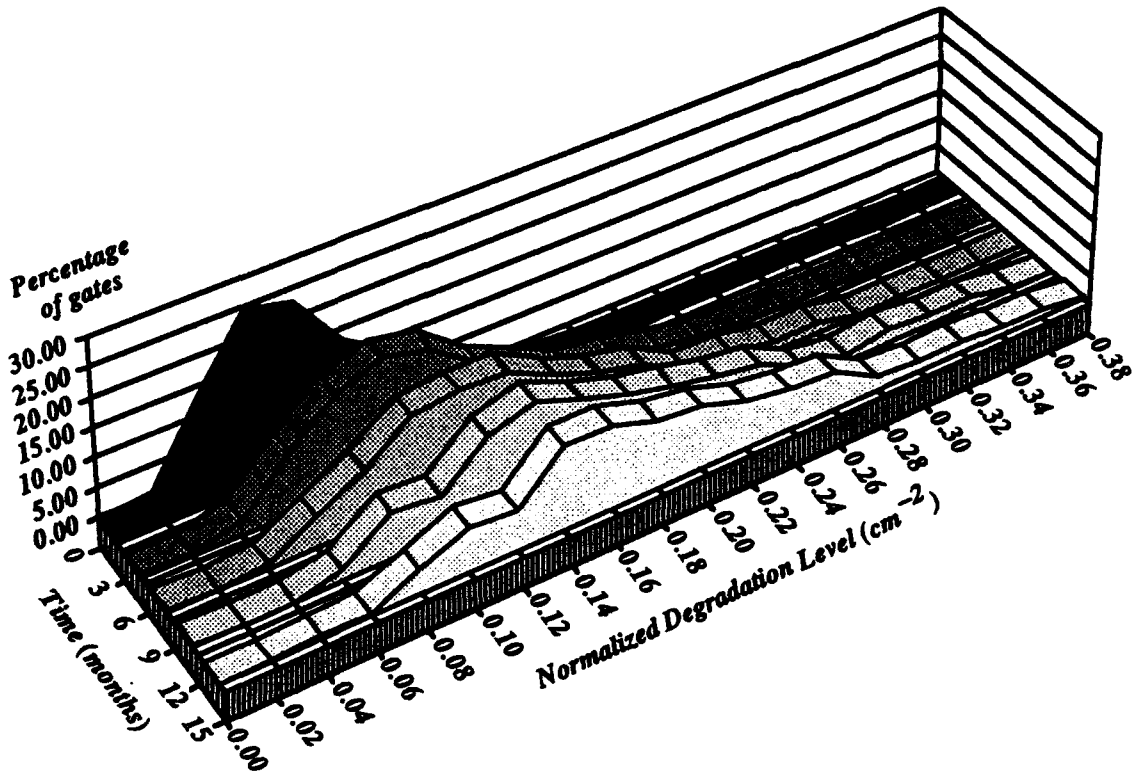


Figure 29: A Degradation Histogram for the 32-bit ALU Circuit

MISSION

OF

ROME LABORATORY

Mission. The mission of Rome Laboratory is to advance the science and technologies of command, control, communications and intelligence and to transition them into systems to meet customer needs. To achieve this, Rome Lab:

a. Conducts vigorous research, development and test programs in all applicable technologies;

b. Transitions technology to current and future systems to improve operational capability, readiness, and supportability;

c. Provides a full range of technical support to Air Force Materiel Command product centers and other Air Force organizations;

d. Promotes transfer of technology to the private sector;

e. Maintains leading edge technological expertise in the areas of surveillance, communications, command and control, intelligence, reliability science, electro-magnetic technology, photonics, signal processing, and computational science.

The thrust areas of technical competence include: Surveillance, Communications, Command and Control, Intelligence, Signal Processing, Computer Science and Technology, Electromagnetic Technology, Photonics and Reliability Sciences.